November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).
Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2006, 21:59:17 ; Search time 93.6 Seconds
(without alignments)
107.136 Million cell updates/sec

Title: US-10-751-235-10
Perfect score: 122
Sequence: 1 LVAEVSEFLFGSGFAIAEGPLWTA 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	122	100.0	24	5	US-10-751-235-10	Sequence 10, Appl
2	122	100.0	222	4	US-10-424-599-265810	Sequence 265810,
3	122	100.0	539	5	US-10-751-235-4	Sequence 4, Appli
4	122	100.0	560	5	US-10-739-930-6317	Sequence 6317, Ap
5	114	93.4	208	5	US-10-751-235-21	Sequence 21, Appl
6	110	90.2	362	5	US-10-751-235-18	Sequence 18, Appl
7	110	90.2	382	4	US-10-425-114-70634	Sequence 70634, A
8	110	90.2	531	4	US-10-437-963-117686	Sequence 117686,
9	110	90.2	545	5	US-10-751-235-17	Sequence 17, Appl
10	110	90.2	561	5	US-10-751-235-16	Sequence 16, Appl
11	110	90.2	588	5	US-10-751-235-52	Sequence 52, Appl
12	110	90.2	624	4	US-10-425-115-332853	Sequence 332853,
13	59	48.4	167	4	US-10-282-122A-52949	Sequence 52949, A
14	59	48.4	342	4	US-10-282-122A-52999	Sequence 52999, A
15	58	47.5	501	4	US-10-425-115-244178	Sequence 244178,
16	58	47.5	588	4	US-10-437-963-139807	Sequence 139807,
17	58	47.5	632	5	US-10-751-235-34	Sequence 34, Appl
18	58	47.5	662	4	US-10-425-114-38827	Sequence 38827, A
19	57	46.7	342	5	US-10-751-235-37	Sequence 37, Appl
20	57	46.7	508	5	US-10-751-235-35	Sequence 35, Appl
21	56	45.9	344	4	US-10-369-493-17393	Sequence 17393, A
22	54	44.3	579	5	US-10-751-235-38	Sequence 38, Appl
23	54	44.3	595	5	US-10-751-235-33	Sequence 33, Appl
24	52.5	43.0	97	3	US-09-738-626-4993	Sequence 4993, Ap
25	51	41.8	1709	4	US-10-156-761-11180	Sequence 11180, A
26	50	41.0	212	4	US-10-424-599-195097	Sequence 195097,
27	50	41.0	350	4	US-10-282-122A-60963	Sequence 60963, A
28	49.5	40.6	287	4	US-10-369-493-17225	Sequence 17225, A
29	49	40.2	148	4	US-10-424-599-223289	Sequence 223289,
30	49	40.2	360	4	US-10-369-493-7853	Sequence 7853, Ap
31	49	40.2	572	4	US-10-167-994-2	Sequence 2, Appli
32	49	40.2	572	4	US-10-017-479-2	Sequence 2, Appli
33	49	40.2	572	5	US-10-718-359-2	Sequence 2, Appli
34	49	40.2	572	6	US-11-097-143-7737	Sequence 7737, Ap
35	49	40.2	572	6	US-11-097-143-27501	Sequence 27501, A
36	48	39.3	107	4	US-10-424-599-273682	Sequence 273682,
37	48	39.3	113	4	US-10-425-115-305171	Sequence 305171,
38	48	39.3	173	4	US-10-282-122A-48087	Sequence 48087, A
39	48	39.3	303	4	US-10-369-493-10381	Sequence 10381, A
40	48	39.3	394	4	US-10-425-114-45724	Sequence 45724, A
41	48	39.3	431	4	US-10-424-599-273679	Sequence 273679,
42	48	39.3	523	4	US-10-230-026-30	Sequence 30, Appl
43	48	39.3	523	5	US-10-486-307-30	Sequence 30, Appl
44	48	39.3	703	4	US-10-437-963-172007	Sequence 172007,
45	48	39.3	1032	4	US-10-389-566-1449	Sequence 1449, Ap

ALIGNMENTS

RESULT 1
US-10-751-235-10
; Sequence 10, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: DellaPenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; TITLE OF INVENTION: Metabolism in Plants
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-751-235-10

Query Match 100.0%; Score 122; DB 5; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.1e-12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LVAEVSEFLFGSGFAIAEGPLWTA 24
||||||||||||||||||||||||
Db 1 LVAEVSEFLFGSGFAIAEGPLWTA 24

RESULT 2
US-10-424-599-265810
; Sequence 265810, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 265810
; LENGTH: 222
; TYPE: PRT

; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_82046C.1.pep
US-10-424-599-265810

Query Match 100.0%; Score 122; DB 4; Length 222;
Best Local Similarity 100.0%; Pred. No. 3.7e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LVAEVSEFLFGSGFAIAEGPLWTA 24
||||||||||||||||||||||||
Db 157 LVAEVSEFLFGSGFAIAEGPLWTA 180

RESULT 3
US-10-751-235-4
; Sequence 4, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: DellaPenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; TITLE OF INVENTION: Metabolism in Plants
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 539
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-751-235-4

Query Match 100.0%; Score 122; DB 5; Length 539;
Best Local Similarity 100.0%; Pred. No. 9.9e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LVAEVSEFLFGSGFAIAEGPLWTA 24
||||||||||||||||||||||||
Db 144 LVAEVSEFLFGSGFAIAEGPLWTA 167

RESULT 4
US-10-739-930-6317
; Sequence 6317, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 6317
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: ARATH-23APR03-C3001_1.p
US-10-739-930-6317

Query Match 100.0%; Score 122; DB 5; Length 560;
Best Local Similarity 100.0%; Pred. No. 1e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LVAEVSEFLFGSGFAIAEGPLWTA 24
||||||||||||||||||||||||
Db 165 LVAEVSEFLFGSGFAIAEGPLWTA 188

RESULT 5
US-10-751-235-21
; Sequence 21, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: DellaPenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; TITLE OF INVENTION: Metabolism in Plants
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Helianthus annuus
US-10-751-235-21

Query Match 93.4%; Score 114; DB 5; Length 208;
Best Local Similarity 95.8%; Pred. No. 6.6e-10;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LVAEVSEFLFGSGFAIAEGPLWTA 24
|||||||||||||||||||| ||||
Db 29 LVAEVSEFLFGSGFAIAEGSLWTA 52

RESULT 6
US-10-751-235-18
; Sequence 18, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: DellaPenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; TITLE OF INVENTION: Metabolism in Plants
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-751-235-18

Query Match 90.2%; Score 110; DB 5; Length 362;
Best Local Similarity 95.7%; Pred. No. 5.4e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LVAEVSEFLFGSGFAIAEGPLWT 23
||||||||||||||||||| |||
Db 141 LVAEVSEFLFGSGFAIAEGALWT 163

RESULT 7
US-10-425-114-70634
; Sequence 70634, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

```
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 70634
;   LENGTH: 382
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: UC-ZMFLMO17009E12_FLI.pep
US-10-425-114-70634

  Query Match             90.2%;  Score 110;  DB 4;  Length 382;
  Best Local Similarity   95.7%;  Pred. No. 5.7e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLWT 23
              |||||||||||||||||||| |||
Db        109 LVAEVSEFLFGSGFAIAEGDLWT 131


RESULT 8
US-10-437-963-117686
; Sequence 117686, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 117686
;   LENGTH: 531
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_21069C.1.pep
US-10-437-963-117686

  Query Match             90.2%;  Score 110;  DB 4;  Length 531;
  Best Local Similarity   95.7%;  Pred. No. 8.3e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLWT 23
              |||||||||||||||||||| |||
Db        152 LVAEVSEFLFGSGFAIAEGALWT 174


RESULT 9
US-10-751-235-17
; Sequence 17, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
;  APPLICANT: DellaPenna, Dean
;  APPLICANT:  Tian, Li
;  APPLICANT:  Kim, Joonyul
;  TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
;  TITLE OF INVENTION:  Metabolism in Plants
;  FILE REFERENCE: MSU-08604
;  CURRENT APPLICATION NUMBER: US/10/751,235
;  CURRENT FILING DATE:  2004-01-02
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
;   LENGTH: 545
;   TYPE: PRT
;   ORGANISM: Hordeum vulgare
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (529)..(529)
;   OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-751-235-17

  Query Match             90.2%;  Score 110;  DB 5;  Length 545;
  Best Local Similarity   95.7%;  Pred. No. 8.5e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLWT 23
              |||||||||||||||||||| |||
Db        136 LVAEVSEFLFGSGFAIAEGALWT 158


RESULT 10
US-10-751-235-16
; Sequence 16, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
;  APPLICANT: DellaPenna, Dean
;  APPLICANT:  Tian, Li
;  APPLICANT:  Kim, Joonyul
;  TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
;  TITLE OF INVENTION:  Metabolism in Plants
;  FILE REFERENCE: MSU-08604
;  CURRENT APPLICATION NUMBER: US/10/751,235
;  CURRENT FILING DATE:  2004-01-02
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
;   LENGTH: 561
;   TYPE: PRT
;   ORGANISM: Oryza sativa
US-10-751-235-16

  Query Match             90.2%;  Score 110;  DB 5;  Length 561;
  Best Local Similarity   95.7%;  Pred. No. 8.8e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLWT 23
              |||||||||||||||||||| |||
Db        152 LVAEVSEFLFGSGFAIAEGALWT 174


RESULT 11
US-10-751-235-52
; Sequence 52, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
;  APPLICANT: DellaPenna, Dean
;  APPLICANT:  Tian, Li
;  APPLICANT:  Kim, Joonyul
;  TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
;  TITLE OF INVENTION:  Metabolism in Plants
;  FILE REFERENCE: MSU-08604
;  CURRENT APPLICATION NUMBER: US/10/751,235
;  CURRENT FILING DATE:  2004-01-02
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 52
;   LENGTH: 588
;   TYPE: PRT
;   ORGANISM: Oryza sativa
US-10-751-235-52

  Query Match             90.2%;  Score 110;  DB 5;  Length 588;
  Best Local Similarity   95.7%;  Pred. No. 9.3e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          1 LVAEVSEFLFGSGFAIAEGPLWT 23
              ||||||||||||||||||| |||
Db        154 LVAEVSEFLFGSGFAIAEGALWT 176


RESULT 12
US-10-425-115-332853
; Sequence 332853, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 332853
;   LENGTH: 624
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(624)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_66678C.1.pep
US-10-425-115-332853

  Query Match             90.2%;  Score 110;  DB 4;  Length 624;
  Best Local Similarity   95.7%;  Pred. No. 9.9e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLWT 23
              ||||||||||||||||||| |||
Db        193 LVAEVSEFLFGSGFAIAEGDLWT 215


RESULT 13
US-10-282-122A-52949
; Sequence 52949, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52949
;   LENGTH: 167
;   TYPE: PRT
;   ORGANISM: Clostridium difficile
US-10-282-122A-52949

  Query Match             48.4%;  Score 59;  DB 4;  Length 167;
  Best Local Similarity   50.0%;  Pred. No. 0.36;
  Matches   10;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGP 20
              :: ||::   | ||:|||||
Db        113 IIDEVTDIFIGMGFSIAEGP 132


RESULT 14
US-10-282-122A-52999
; Sequence 52999, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52999
;   LENGTH: 342
```

```
;   TYPE: PRT
;   ORGANISM: Clostridium difficile
US-10-282-122A-52999

  Query Match             48.4%;  Score 59;  DB 4;  Length 342;
  Best Local Similarity   50.0%;  Pred. No. 0.8;
  Matches   10;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGP 20
              :: ||::   | ||:|||||
Db        116 IIDEVTDIFIGMGFSIAEGP 135


RESULT 15
US-10-425-115-244178
; Sequence 244178, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 244178
;   LENGTH: 501
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_154276C.1.pep
US-10-425-115-244178

  Query Match             47.5%;  Score 58;  DB 4;  Length 501;
  Best Local Similarity   40.9%;  Pred. No. 1.8;
  Matches    9;  Conservative    7;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLW 22
              ::||: ||: |:|   |:| :|
Db         64 ILAEILEFVMGTGLIPADGEIW 85


Search completed: March  7, 2006, 22:04:14
Job time : 94.6 secs
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2006, 21:43:03 ; Search time 352.2 Seconds
(without alignments)
44.911 Million cell updates/sec

Title: US-10-751-235-11
Perfect score: 177
Sequence: 1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	177	100.0	36	9	AEB16913	Aeb16913 Thale cre
2	177	100.0	539	9	AEB16908	Aeb16908 Thale cre
3	177	100.0	560	8	ADT56240	Adt56240 Plant pol
4	75.5	42.7	386	3	AAY44631	Aay44631 A. thalia
5	75.5	42.7	386	4	AAB85456	Aab85456 Arabidops
6	63	35.6	260	8	ADM48270	Adm48270 Polypepti
7	62.5	35.3	2016	4	ABB63911	Abb63911 Drosophil
8	60	33.9	222	3	AAG26703	Aag26703 Arabidops
9	60	33.9	355	5	ABB93995	Abb93995 Herbicida
10	60	33.9	389	3	AAG54022	Aag54022 Arabidops
11	60	33.9	389	3	AAG52506	Aag52506 Arabidops
12	59.5	33.6	317	3	AAG59958	Aag59958 Arabidops
13	59.5	33.6	317	3	AAG10212	Aag10212 Arabidops
14	59.5	33.6	319	3	AAG10211	Aag10211 Arabidops
15	59.5	33.6	319	3	AAG59957	Aag59957 Arabidops
16	59.5	33.6	319	5	ABB93989	Abb93989 Herbicida
17	59.5	33.6	319	8	ADT55654	Adt55654 Plant pol
18	59	33.3	1545	4	ABB65577	Abb65577 Drosophil
19	57	32.2	100	3	AAG43792	Aag43792 Arabidops
20	56	31.6	48	4	ABG04863	Abg04863 Novel hum
21	56	31.6	407	3	AAG22122	Aag22122 Arabidops
22	56	31.6	434	3	AAG22121	Aag22121 Arabidops
23	56	31.6	474	8	ADN74381	Adn74381 Thale cre
24	56	31.6	883	4	ABB62230	Abb62230 Drosophil
25	55.5	31.4	370	8	ADX66893	Adx66893 Plant ful
26	55	31.1	137	3	AAG17405	Aag17405 Arabidops
27	55	31.1	143	3	AAG17404	Aag17404 Arabidops
28	54.5	30.8	92	3	AAG16401	Aag16401 Arabidops
29	54.5	30.8	121	3	AAG16400	Aag16400 Arabidops
30	54.5	30.8	416	8	ADI42851	Adi42851 Plant tra
31	54.5	30.8	416	8	ADO03046	Ado03046 Thalecres
32	54.5	30.8	898	7	ABM87814	Abm87814 Rice abio
33	54.5	30.8	2027	4	AAE10987	Aae10987 Mouse pre
34	54	30.5	205	5	ABB79956	Abb79956 Arabidops
35	54	30.5	205	6	ABR40828	Abr40828 Arabidops
36	54	30.5	251	8	ADT60146	Adt60146 Plant pol
37	54	30.5	271	7	ABM86366	Abm86366 Rice abio
38	54	30.5	307	4	ABG22889	Abg22889 Novel hum
39	54	30.5	403	3	AAG53830	Aag53830 Arabidops
40	54	30.5	403	3	AAG16661	Aag16661 Arabidops
41	54	30.5	430	5	AAU93039	Aau93039 Arabidops
42	54	30.5	430	5	ABB79954	Abb79954 Arabidops
43	54	30.5	430	5	ABB79955	Abb79955 Arabidops
44	54	30.5	430	7	ADB31801	Adb31801 Plant (A.
45	54	30.5	430	7	ADC46705	Adc46705 Thalecres

ALIGNMENTS

RESULT 1
AEB16913
ID AEB16913 standard; peptide; 36 AA.
XX
AC AEB16913;
XX
DT 08-SEP-2005 (first entry)
XX
DE Thale cress chloroplast-targeting N-terminal peptide, SEQ ID NO: 11.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant;
KW cytochrome P450.
XX
OS Arabidopsis thaliana.
XX
PN US2005150002-A1.
XX
PD 07-JUL-2005.
XX
PF 02-JAN-2004; 2004US-00751235.
XX
PR 02-JAN-2004; 2004US-00751235.
XX
PA (DELL/) DELLAPENNA D.
PA (TIAN/) TIAN L.
PA (KIMJ/) KIM J.
XX
PI Dellapenna D, Tian L, Kim J;
XX
DR WPI; 2005-487984/49.
XX
PT New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.
XX
PS Claim 8; SEQ ID NO 11; 135pp; English.
XX
CC The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is thale cress chlorophyll-targeting N
CC -terminal transit peptide.
XX

```
SQ   Sequence 36 AA;

  Query Match             100.0%;  Score 177;  DB 9;  Length 36;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-17;
  Matches   36;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36
              |||||||||||||||||||||||||||||||||||
Db          1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36


RESULT 2
AEB16908
ID   AEB16908 standard; protein; 539 AA.
XX
AC   AEB16908;
XX
DT   08-SEP-2005  (first entry)
XX
DE   Thale cress LUT1 (cytochrome P450 97C1) protein, SEQ ID NO: 4.
XX
KW   Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C1;
KW   cytochrome P450 97C1.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .36
FT                   /note= "Chloroplast-targeting N-terminal transit peptide"
FT   Misc-difference 166. .167
FT                   /note= "Encoded by bases of positions 496 to 650"
FT   Misc-difference 265. .266
FT                   /note= "Encoded by bases of positions 942 to 1037"
FT   Misc-difference 332. .333
FT                   /note= "Encoded by bases of positions 1242 to 1326"
FT   Region          347. .352
FT                   /note= "Molecular oxygen binding pocket"
FT   Misc-difference 365. .366
FT                   /note= "Encoded by bases of positions 1420 to 1523"
FT   Misc-difference 413. .414
FT                   /note= "Encoded by bases of positions 1662 to 1841"
FT   Misc-difference 449. .450
FT                   /note= "Encoded by bases of positions 1839 to 12005"
FT   Misc-difference 476. .477
FT                   /note= "Encoded by bases of positions 2081 to 2183"
FT   Region          480. .489
FT                   /note= "Cysteine motif region"
FT   Misc-difference 530. .531
FT                   /note= "Encoded by bases of positions 2340 to 2440"
XX
PN   US2005150002-A1.
XX
PD   07-JUL-2005.
XX
PF   02-JAN-2004; 2004US-00751235.
XX
PR   02-JAN-2004; 2004US-00751235.
XX
PA   (DELL/) DELLAPENNA D.
PA   (TIAN/) TIAN L.
PA   (KIMJ/) KIM J.
XX
PI   Dellapenna D,  Tian L,  Kim J;
XX
DR   WPI; 2005-487984/49.
XX
PT   New expression vector comprising a nucleic acid sequence encoding a
PT   polypeptide having monooxygenase P450 activity, useful in altering the
PT   carotenoid production in a plant for enhancing production of specific
PT   carotenoid compounds.
XX
PS   Claim 9; SEQ ID NO 4; 135pp; English.
XX
CC   The present invention relates to genes, proteins and methods comprising
CC   carotenoid monooxygenases in the cytochrome P450 family. The invention
CC   also relates to altering carotenoid ratios in plants and microorganisms
CC   using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC   invention is useful in altering the carotenoid production in a plant for
CC   enhancing production of specific carotenoid compounds that are potent
CC   antioxidants. The present sequence is thale cress LUT1 [cytochrome P450
CC   monooxygenase (CYP97C1); At3g53130 gene] protein. Note: The current
CC   sequence is that of thale cress LUT1 protein that is encoded by cDNA
CC   located on chromosome 3.
XX
SQ   Sequence 539 AA;

  Query Match             100.0%;  Score 177;  DB 9;  Length 539;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-16;
  Matches   36;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36
              |||||||||||||||||||||||||||||||||||
Db          1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36


RESULT 3
ADT56240
ID   ADT56240 standard; protein; 560 AA.
XX
AC   ADT56240;
XX
DT   13-JAN-2005  (first entry)
XX
DE   Plant polypeptide, SEQ ID 6317.
XX
KW   Plant; transgenic; cold tolerance; growth rate; drought tolerance;
KW   disease resistance; galactomannan production; plant growth regulator;
KW   heat tolerance; herbicide tolerance; lignin production;
KW   extreme osmotic condition tolerance; pathogens resistance;
KW   pest resistance; yield improvement; seed oil yield; seed protein yield.
XX
OS   Viridiplantae.
XX
PN   US2004216190-A1.
XX
PD   28-OCT-2004.
XX
PF   18-DEC-2003; 2003US-00739930.
XX
PR   28-APR-2003; 2003US-00424599.
PR   28-APR-2003; 2003US-00425115.
XX
PA   (KOVA/) KOVALIC D K.
XX
PI   Kovalic DK;
XX
DR   WPI; 2004-757369/74.
XX
PT   New recombinant DNA constructs useful in the field of biochemistry and
PT   genetics, and in particular for producing transgenic plants with improved
PT   biological characteristics.
XX
PS   Claim 2; SEQ ID NO 6317; 14pp; English.
XX
CC   The invention relates a recombinant DNA construct comprising a
CC   polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
CC   1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC   (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
CC   Arabidopsis, wheat and rape but the specification does not indicate which
CC   sequences is derived from which organism. Also included is a method of
CC   producing a plant having an improved property, comprising transforming a
CC   plant with a recombinant DNA construct comprising a promoter region
CC   functional in a plant cell operably joined to a polynucleotide encoding a
CC   polypeptide associated with the property, and growing the transformed
CC   plant. The property is selected from improving plant cold tolerance, for
CC   manipulating growth rate in plant cells by modification of the cell cycle
```

CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant disease, for galactomannan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and
CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to extreme osmotic and drought conditions, and improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant protein
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040216190.
XX
SQ Sequence 560 AA;

Query Match 100.0%; Score 177; DB 8; Length 560;
Best Local Similarity 100.0%; Pred. No. 4.7e-16;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36
||||||||||||||||||||||||||||||||||||
Db 22 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 57

RESULT 4
AAY44631
ID AAY44631 standard; protein; 386 AA.
XX
AC AAY44631;
XX
DT 07-APR-2000 (first entry)
XX
DE A. thaliana AIR synthetase pre-protein.
XX
KW 5'-phosphoribosyl-5-aminoimidazole synthetase; weed growth;
KW purine biosynthesis; recombinant AIR synthetase; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO9967402-A2.
XX
PD 29-DEC-1999.
XX
PF 22-JUN-1999; 99WO-EP004309.
XX
PR 24-JUN-1998; 98US-00103895.
XX
PA (NOVS) NOVARTIS AG.
PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX
PI Ward ER, Guyer CD, Potter SL, Subramanian V, Walters E;
XX
DR WPI; 2000-136986/12.
DR N-PSDB; AAZ49679.
XX
PT New isolated enzyme with phosphoribosyl-5-aminoimidazole synthetase
PT activity, used to screen for herbicides or to produce herbicide-tolerant
PT plants.
XX
PS Claim 3; Page 53-54; 58pp; English.
XX
CC The present sequence is A. thaliana AIR (5'-phosphoribosyl-5- -
CC aminoimidazole) synthetase pre-protein, an enzyme involved in de novo
CC purine biosynthesis. The cDNA is deposited in E. coli strain DH5apASM
CC designated as NRRL accession number B-21976. Recombinantly produced AIR
CC synthetase enzyme can be used in screening chemicals for herbicidal
CC activity. Plants having modified enzyme activity are tolerant to
CC inhibition by herbicides. Herbicide-tolerant AIR synthetase is also
CC useful as a selectable marker in plants
XX
SQ Sequence 386 AA;

Query Match 42.7%; Score 75.5; DB 3; Length 386;
Best Local Similarity 48.6%; Pred. No. 0.063;
Matches 17; Conservative 5; Mismatches 10; Indels 3; Gaps 1;

Qy 1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSI 35
||: : ||| ||||:| :| |||| ||:
Db 1 MEARILQSSSSCYSSLYTVNRSRFSSPKP---FSV 32

RESULT 5
AAB85456
ID AAB85456 standard; protein; 386 AA.
XX
AC AAB85456;
XX
DT 25-SEP-2001 (first entry)
XX
DE Arabidopsis AIR synthetase.
XX
KW 5'-phosphoribosyl-5-aminoimidazole synthetase; AIR synthetase;
KW Arabidopsis; transgenic; herbicide; enzyme.
XX
OS Arabidopsis thaliana.
XX
PN US6271445-B1.
XX
PD 07-AUG-2001.
XX
PF 22-JUN-1999; 99US-00338420.
XX
PR 24-JUN-1998; 98US-0155234P.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Ward ER, Guyer DC, Potter SL, Subramanian V, Walters E;
XX
DR WPI; 2001-475191/51.
DR N-PSDB; AAH23384.
XX
PT Novel isolated nucleic acid encoding a 5'-phosphoribosyl-5-aminoimidazole
PT synthase, used to transform plant cells, for use in screening for
PT herbicides.
XX
PS Claim 1; Col 29-32; 21pp; English.
XX
CC The invention relates to a 5'-phosphoribosyl-5-aminoimidazole (AIR)
CC synthetase from Arabidopsis. The AIR synthetase protein can be expressed
CC by standard recombinant methodology. Nucleic acids encoding the AIR
CC synthetase are used for producing transgenic plants and plant cells, for
CC screening herbicidal compounds which inhibit the enzymatic activity of
CC AIR synthetase. The present sequence represents an Arabidopsis AIR
CC synthetase sequence
XX
SQ Sequence 386 AA;

Query Match 42.7%; Score 75.5; DB 4; Length 386;
Best Local Similarity 48.6%; Pred. No. 0.063;
Matches 17; Conservative 5; Mismatches 10; Indels 3; Gaps 1;

Qy 1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSI 35
||: : ||| ||||:| :| |||| ||:

```
Db          1 MEARILQSSSSCYSSLYTVNRSRFSSPKP---FSV 32


RESULT 6
ADM48270
ID   ADM48270 standard; protein; 260 AA.
XX
AC   ADM48270;
XX
DT   03-JUN-2004  (first entry)
XX
DE   Polypeptide sequence #320 useful in producing transgenic plants.
XX
KW   Plant; maize; rice; sorghum; thale cress; soybean; wheat; transgenic;
KW   osmotic stress; sugar transport; cell cycle pathway; plant height;
KW   carbohydrate transport; crop productivity; plant growth;
KW   stress resistance; disease resistance; insect resistance; heat tolerance;
KW   nitrogen assimilation; water stress tolerance;
KW   photosynthetic carbon fixation; virus resistance; gene therapy.
XX
OS   Glycine max.
XX
PN   US2003233670-A1.
XX
PD   18-DEC-2003.
XX
PF   04-DEC-2002; 2002US-00310154.
XX
PR   04-DEC-2001; 2001US-0337358P.
XX
PA   (EDGE/) EDGERTON M D.
PA   (CHOM/) CHOMET P S.
PA   (LACC/) LACCETTI L B.
XX
PI   Edgerton MD,  Chomet PS,  Laccetti LB;
XX
DR   WPI; 2004-061374/06.
DR   N-PSDB; ADM47902.
XX
PT   New polynucleotide, useful for manipulating plant protein quality,
PT   improving plant growth, yield and crop productivity or grain composition
PT   or producing plants with improved properties.
XX
PS   Claim 8; SEQ ID NO 688; 144pp; English.
XX
CC   The present invention relates to polynucleotide sequences, and the
CC   proteins they encode. The sequences are isolated from a variety of
CC   organisms such as plants (e.g. maize, rice, sorghum, thale cress,
CC   soybean, and wheat), cyanobacteria, bacteria, yeast and other fungi. The
CC   polynucleotide and polypeptide sequences of the invention are useful in
CC   the production of transgenic plants that have improved properties. Also
CC   disclosed are methods of producing fertile transgenic plants, preferably
CC   maize, with desired phenotypes. The polynucleotide and polypeptide
CC   sequences are useful for improving plants by providing protection against
CC   osmotic stress, improving altering sugar transport and/or metabolism,
CC   modifying the cell cycle pathway, reducing plant height, modifying
CC   carbohydrate transport, improving crop productivity, improving plant
CC   growth and stress resistance, improving disease resistance, improving
CC   insect resistance, improving cold or heat tolerance, improving nitrogen
CC   assimilation, improving stalk strength, improving water stress tolerance,
CC   improving photosynthetic carbon fixation, improving biotic and abiotic
CC   stress resistance, improving resistance to oxidative stress, providing
CC   increased vigour, reducing senescence, and conferring virus resistance.
CC   The present sequence represents a polypeptide sequence of the invention.
CC   Note: The sequence data for this patent is not provided in the printed
CC   specification but is obtained in electronic format from the USPTO website
CC   at seqdata.uspto.gov.
XX
SQ   Sequence 260 AA;

  Query Match             35.6%;  Score 63;  DB 8;  Length 260;
  Best Local Similarity   46.2%;  Pred. No. 2.3;
  Matches   12;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         10 SSSYSSLFTAKPTRLLSPKPKFTFSI 35
              |:: :|||:| || | ||||  :  :
Db          4 SATSASLFSANPTPLFSPKPSLSLHL 29


RESULT 7
ABB63911
ID   ABB63911 standard; protein; 2016 AA.
XX
AC   ABB63911;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 18525.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL08014.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Disclosure; SEQ ID NO 18525; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 2016 AA;

  Query Match             35.3%;  Score 62.5;  DB 4;  Length 2016;
  Best Local Similarity   47.4%;  Pred. No. 30;
  Matches   18;  Conservative    4;  Mismatches   11;  Indels    5;  Gaps    2;

Qy          3 SSLFSPSS-SSYSSLFTAKP----TRLLSPKPKFTFSI 35
              |||||||   | ||||   |    | :|:| |  | ::
Db        986 SSLFSPSQVDSTSSLFAPPPTSANTTILTPAPDTTSTL 1023


RESULT 8
AAG26703
ID   AAG26703 standard; protein; 222 AA.
XX
AC   AAG26703;
XX
DT   17-OCT-2000  (first entry)
XX
```

DE Arabidopsis thaliana protein fragment SEQ ID NO: 31259.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.

PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 33.9%; Score 60; DB 3; Length 222;
Best Local Similarity 40.0%; Pred. No. 5;
Matches 14; Conservative 5; Mismatches 16; Indels 0; Gaps 0;

```
Qy          1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSI 35
              ||: :   ||| ||||:|   :|  |      ||:
Db          1 MEARILQSSSSCYSSLYTVNRSRFSSVSSPKLFSV 35
```

RESULT 9
ABB93995
ID ABB93995 standard; protein; 355 AA.
XX
AC ABB93995;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 3206.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP009892.
XX
PR 28-AUG-2001; 2001WO-EP009892.
XX
PA (FARB) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms.
XX
PS Claim 5; SEQ ID NO 3206; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins (ABB90790-ABB94016)
CC for herbicidally active compounds, comprising aligning and comparing
CC nucleic acid or amino acid sequences from plant with nucleic acid or
CC amino acid sequences from non-plant organisms using suitable search
CC parameters, where plant sequences having an E-value greater by a factor
CC of 3 than the E-value of most similar non-plant sequences are selected.
CC The polypeptides or nucleic acids encoding them are useful for
CC identifying modulators. The identified modulators are useful as
CC herbicides
XX
SQ Sequence 355 AA;

Query Match 33.9%; Score 60; DB 5; Length 355;
Best Local Similarity 42.1%; Pred. No. 8.7;
Matches 16; Conservative 6; Mismatches 10; Indels 6; Gaps 2;

```
Qy          1 MESSLFSPSSSSYSSLFTAKPTR---LLSPKPKFTFSI 35
              ||: :   ||| ||||:    :|   : ||||   ||:
Db          1 MEARILQSSSSCYSSLYAVNRSRFSSVSSPKP---FSV 35
```

RESULT 10
AAG54022
ID AAG54022 standard; protein; 389 AA.
XX
AC AAG54022;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 68835.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.

PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 33.9%; Score 60; DB 3; Length 389;
Best Local Similarity 42.1%; Pred. No. 9.7;
Matches 16; Conservative 6; Mismatches 10; Indels 6; Gaps 2;

```
Qy          1 MESSLFSPSSSSYSSLFTAKPTR---LLSPKPKFTFSI 35
              ||: :   ||| ||||:    :|   : ||||   ||:
Db          1 MEARILQSSSSCYSSLYAVNRSRFSSVSSPKP---FSV 35
```

RESULT 11
AAG52506
ID AAG52506 standard; protein; 389 AA.
XX
AC AAG52506;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 66753.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.

PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 33.9%; Score 60; DB 3; Length 389;
Best Local Similarity 42.1%; Pred. No. 9.7;
Matches 16; Conservative 6; Mismatches 10; Indels 6; Gaps 2;

```
Qy          1 MESSLFSPSSSSYSSLFTAKPTR---LLSPKPKFTFSI 35
              ||: :   ||| ||||:    :|   : ||||   ||:
Db          1 MEARILQSSSSCYSSLYAVNRSRFSSVSSPKP---FSV 35
```

RESULT 12
AAG59958
ID AAG59958 standard; protein; 317 AA.
XX
AC AAG59958;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 77612.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.

PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 33.6%; Score 59.5; DB 3; Length 317;
Best Local Similarity 43.8%; Pred. No. 9;
Matches 14; Conservative 7; Mismatches 8; Indels 3; Gaps 1;

```
Qy          3 SSLFSPSSSSYSSLFTAKPTRL---LSPKPKF 31
              ::|:| || || |  |:||:|   .| |: :|
Db          5 NTLYSSSSPSYLSPLTSKPSRFSKNLRPRAQF 36
```

RESULT 13
AAG10212
ID AAG10212 standard; protein; 317 AA.
XX
AC AAG10212;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 8440.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.

PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 33.6%; Score 59.5; DB 3; Length 317;
Best Local Similarity 43.8%; Pred. No. 9;
Matches 14; Conservative 7; Mismatches 8; Indels 3; Gaps 1;

```
Qy          3 SSLFSPSSSSYSSLFTAKPTRL---LSPKPKF 31
              ::|:| || || |  |:||:|    | |: :|
Db          5 NTLYSSSSPSYLSPLTSKPSRFSKNLRPRAQF 36
```

RESULT 14
AAG10211
ID AAG10211 standard; protein; 319 AA.
XX
AC AAG10211;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 8439.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.

PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 33.6%; Score 59.5; DB 3; Length 319;
Best Local Similarity 43.8%; Pred. No. 9;
Matches 14; Conservative 7; Mismatches 8; Indels 3; Gaps 1;

```
Qy          3 SSLFSPSSSSYSSLFTAKPTRL---LSPKPKF 31
              ::|:| || || |  |:||:|    | |: :|
Db          7 NTLYSSSSPSYLSPLTSKPSRFSKNLRPRAQF 38
```

RESULT 15
AAG59957
ID AAG59957 standard; protein; 319 AA.
XX
AC AAG59957;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 77611.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.

PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 33.6%; Score 59.5; DB 3; Length 319;
Best Local Similarity 43.8%; Pred. No. 9;
Matches 14; Conservative 7; Mismatches 8; Indels 3; Gaps 1;

```
Qy          3 SSLFSPSSSSYSSLFTAKPTRL---LSPKPKF 31
              ::|:| || || |  |:||:|    | |: :|
Db          7 NTLYSSSSPSYLSPLTSKPSRFSKNLRPRAQF 38
```

Search completed: March 7, 2006, 21:53:10

Job time : 356.2 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2006, 21:58:31 ; Search time 40.2 Seconds
(without alignments)
74.038 Million cell updates/sec

Title: US-10-751-235-11
Perfect score: 177
Sequence: 1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	75.5	42.7	386	2	US-09-338-420-2	Sequence 2, Appli
2	59.5	33.6	319	2	US-10-159-901-57	Sequence 57, Appl
3	54.5	30.8	159	2	US-09-248-796A-21631	Sequence 21631, A
4	54	30.5	117	2	US-09-248-796A-23978	Sequence 23978, A
5	53.5	30.2	169	2	US-09-248-796A-21929	Sequence 21929, A
6	53.5	30.2	1564	2	US-09-487-558B-244	Sequence 244, App
7	53	29.9	829	1	US-08-346-455B-34	Sequence 34, Appl
8	53	29.9	829	2	US-08-977-221-34	Sequence 34, Appl
9	53	29.9	829	2	US-09-483-831B-34	Sequence 34, Appl
10	53	29.9	829	4	PCT-US95-06613-34	Sequence 34, Appl
11	53	29.9	915	1	US-08-346-455B-69	Sequence 69, Appl
12	53	29.9	915	2	US-08-977-221-69	Sequence 69, Appl
13	53	29.9	915	2	US-09-483-831B-69	Sequence 69, Appl
14	53	29.9	915	4	PCT-US95-06613-69	Sequence 69, Appl
15	53	29.9	1971	2	US-09-914-272A-1	Sequence 1, Appli
16	53	29.9	1971	2	US-10-638-333-1	Sequence 1, Appli
17	53	29.9	1971	2	US-10-747-133A-1	Sequence 1, Appli
18	52.5	29.7	288	2	US-09-270-767-42632	Sequence 42632, A
19	51	28.8	72	2	US-09-107-433-2902	Sequence 2902, Ap
20	51	28.8	119	2	US-09-270-767-35788	Sequence 35788, A
21	51	28.8	119	2	US-09-270-767-51005	Sequence 51005, A
22	50	28.2	688	2	US-09-489-039A-9153	Sequence 9153, Ap
23	49.5	28.0	680	1	US-08-542-363-4	Sequence 4, Appli
24	49.5	28.0	680	2	US-09-100-089-4	Sequence 4, Appli
25	49.5	28.0	680	2	US-09-670-827-4	Sequence 4, Appli
26	49.5	28.0	680	2	US-09-827-949-4	Sequence 4, Appli
27	49.5	28.0	1620	1	US-08-542-363-2	Sequence 2, Appli
28	49.5	28.0	1620	2	US-09-100-089-2	Sequence 2, Appli
29	49.5	28.0	1620	2	US-09-670-827-2	Sequence 2, Appli
30	49.5	28.0	1620	2	US-09-827-949-2	Sequence 2, Appli
31	49	27.7	53	2	US-08-905-223-381	Sequence 381, App
32	49	27.7	140	2	US-09-673-395A-299	Sequence 299, App
33	49	27.7	225	2	US-09-252-991A-18792	Sequence 18792, A
34	49	27.7	286	2	US-09-248-796A-22818	Sequence 22818, A
35	49	27.7	334	2	US-10-012-231A-41	Sequence 41, Appl
36	49	27.7	334	2	US-10-015-389A-41	Sequence 41, Appl
37	49	27.7	334	2	US-10-006-768A-41	Sequence 41, Appl
38	49	27.7	334	2	US-10-015-671A-41	Sequence 41, Appl
39	49	27.7	334	2	US-10-015-393A-41	Sequence 41, Appl
40	49	27.7	334	2	US-10-011-833A-41	Sequence 41, Appl
41	49	27.7	334	2	US-10-006-041A-41	Sequence 41, Appl
42	49	27.7	334	2	US-10-012-064A-41	Sequence 41, Appl
43	49	27.7	1013	2	US-09-527-084A-2	Sequence 2, Appli
44	49	27.7	2972	2	US-09-579-181-2	Sequence 2, Appli
45	49	27.7	3118	2	US-09-579-181-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-338-420-2
; Sequence 2, Application US/09338420
; Patent No. 6271445
; GENERAL INFORMATION:
; APPLICANT: Ward, Eric R
; APPLICANT: Guyer, David
; APPLICANT: Potter, Sharon
; APPLICANT: Subramanian, Mani
; APPLICANT: Walters, Eric
; TITLE OF INVENTION: Methods to screen herbicidal compounds
; TITLE OF INVENTION: and uses thereof
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6271445artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6271445th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/338,420
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/103,895
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Stults, Larry W
; REGISTRATION NUMBER: 34,025
; REFERENCE/DOCKET NUMBER: CGC1999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8666
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 386 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-338-420-2

Query Match 42.7%; Score 75.5; DB 2; Length 386;
Best Local Similarity 48.6%; Pred. No. 0.017;
Matches 17; Conservative 5; Mismatches 10; Indels 3; Gaps 1;

```
Qy          1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSI 35
              ||: :   ||| ||||:|   :|  ||||   ||:
Db          1 MEARILQSSSSCYSSLYTVNRSRFSSPKP---FSV 32


RESULT 2
US-10-159-901-57
; Sequence 57, Application US/10159901
; Patent No. 6887688
; GENERAL INFORMATION:
;  APPLICANT: LAGARIAS, JOHN
;  APPLICANT:  KOICHI, TAKAYUKI
;  APPLICANT:  FRANKENBERG, NICOLE
;  APPLICANT:  GAMBETTA, GREGORY
;  APPLICANT:  MONTGOMERY, BERONDA
;  TITLE OF INVENTION: LIGHT CONTROLLED GENE EXPRESSION UTILIZING HETEROLOGOUS PHYTOCHRO
;  FILE REFERENCE: 407T-907731US
;  CURRENT APPLICATION NUMBER: US/10/159,901
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: 60/294,463
;  PRIOR FILING DATE: 2001-05-29
;  NUMBER OF SEQ ID NOS: 57
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 57
;   LENGTH: 319
;   TYPE: PRT
;   ORGANISM: Arapidopsis thaliana
US-10-159-901-57

  Query Match             33.6%;  Score 59.5;  DB 2;  Length 319;
  Best Local Similarity   43.8%;  Pred. No. 2.3;
  Matches   14;  Conservative    7;  Mismatches    8;  Indels    3;  Gaps    1;

Qy          3 SSLFSPSSSSYSSLFTAKPTRL---LSPKPKF 31
              ::|:| || || |  |:||:|    | |: :|
Db          7 NTLYSSSSPSYLSPLTSKPSRFSKNLRPRAQF 38


RESULT 3
US-09-248-796A-21631
; Sequence 21631, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 21631
;   LENGTH: 159
;   TYPE: PRT
;   ORGANISM: Candida albicans
US-09-248-796A-21631

  Query Match             30.8%;  Score 54.5;  DB 2;  Length 159;
  Best Local Similarity   37.8%;  Pred. No. 5.1;
  Matches   14;  Conservative    6;  Mismatches   10;  Indels    7;  Gaps    1;

Qy          3 SSLFSPSSSSYSSLFTAK-------PTRLLSPKPKFT 32
              |   ||| || ::  ||:       || :::| || |
Db         17 SKKISPSLSSITTTTTAEDESINDLPTNVITPTPKTT 53


RESULT 4
US-09-248-796A-23978
; Sequence 23978, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 23978
;   LENGTH: 117
;   TYPE: PRT
;   ORGANISM: Candida albicans
US-09-248-796A-23978

  Query Match             30.5%;  Score 54;  DB 2;  Length 117;
  Best Local Similarity   41.0%;  Pred. No. 4.1;
  Matches   16;  Conservative    5;  Mismatches   12;  Indels    6;  Gaps    2;

Qy          3 SSLFSPSSSSYSSL-FTAKPTRLL-----SPKPKFTFSI 35
              | |    || || : ||  |||:|     : ||| |: :
Db         25 SFLSQQHSSKYSYIHFTMIPTRILLNGAKNVKPKLTYPV 63


RESULT 5
US-09-248-796A-21929
; Sequence 21929, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 21929
;   LENGTH: 169
;   TYPE: PRT
;   ORGANISM: Candida albicans
US-09-248-796A-21929

  Query Match             30.2%;  Score 53.5;  DB 2;  Length 169;
  Best Local Similarity   34.1%;  Pred. No. 7.6;
  Matches   15;  Conservative    9;  Mismatches    7;  Indels   13;  Gaps    2;

Qy          2 ESSLF------SPSSSSYSSLFTAK-------PTRLLSPKPKFT 32
              :||::      ||| || ::  ||:       || :::| || |
Db         25 KSSIYLTQKRLSPSLSSITTTTTAEDESINDLPTNVITPTPKTT 68


RESULT 6
US-09-487-558B-244
; Sequence 244, Application US/09487558B
; Patent No. 6949356
; GENERAL INFORMATION:
;  APPLICANT: Busby, Robert
;  APPLICANT:  Cali, Brian
;  APPLICANT:  Hecht, Peter
;  APPLICANT:  Holtzman, Doug
;  APPLICANT:  Madden, Kevin
;  APPLICANT:  Maxon, Mary
;  APPLICANT:  Milne, Todd
;  APPLICANT:  No. 6949356man, Thea
;  APPLICANT:  Royer, John
```

```
;  APPLICANT:  Salama, Sofie
;  APPLICANT:  Sherman, Amir
;  APPLICANT:  Silva, Jeff
;  APPLICANT:  Summers, Eric
;  TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
;  FILE REFERENCE: 109272.130
;  CURRENT APPLICATION NUMBER: US/09/487,558B
;  CURRENT FILING DATE:  2000-01-19
;  PRIOR APPLICATION NUMBER: US 60/487,558
;  PRIOR FILING DATE: 1999-10-20
;  NUMBER OF SEQ ID NOS: 446
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 244
;   LENGTH: 1564
;   TYPE: PRT
;   ORGANISM: Saccharomyces cerevisiae
US-09-487-558B-244

  Query Match             30.2%;  Score 53.5;  DB 2;  Length 1564;
  Best Local Similarity   37.5%;  Pred. No. 1.1e+02;
  Matches   12;  Conservative    9;  Mismatches    8;  Indels    3;  Gaps    1;

Qy          4 SLFSPSSSSYSSLFTAKPTRLLSPKPKFTPSI 35
              ||: ||: :::| |:  ||:| :     ||:|
Db        612 SLYRPSADAPASTFSDVPTKLAT---AVTFNI 640


RESULT 7
US-08-346-455B-34
; Sequence 34, Application US/08346455B
; Patent No. 5731167
;  GENERAL INFORMATION:
;    APPLICANT:  UNITED STATES OF AMERICA; DEPT.
;    APPLICANT:  OF HEALTH AND HUMAN SERVICES
;    TITLE OF INVENTION:  MOTILITY STIMULATING
;    TITLE OF INVENTION:  PROTEIN USEFUL IN CANCER DIAGNOSIS AND
;    TITLE OF INVENTION:  THERAPY
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  U.S.A.
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/346,455B
;      FILING DATE:  28-NOV-1994
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/06613
;      FILING DATE:  24-MAY-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/249,182
;      FILING DATE:  25-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/822,043
;      FILING DATE:  17-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DOROTHY R. AUTH
;      REGISTRATION NUMBER:  36,434
;      REFERENCE/DOCKET NUMBER:  2026-4149PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  829
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  Unknown
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  No
;    ORIGINAL SOURCE:
;      ORGANISM:  Human
;      STRAIN:
;      INDIVIDUAL ISOLATE:
;      DEVELOPMENTAL STAGE:
;      HAPLOTYPE:
;      TISSUE TYPE:
;      CELL TYPE:  Melanoma
;      CELL LINE:  A2058
;      ORGANELLE:
;    FEATURE:
;      NAME/KEY:
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:  Putative protein
;      OTHER INFORMATION:  sequence of A2058 Autotaxin
US-08-346-455B-34

  Query Match             29.9%;  Score 53;  DB 1;  Length 829;
  Best Local Similarity   48.0%;  Pred. No. 59;
  Matches   12;  Conservative    3;  Mismatches    8;  Indels    2;  Gaps    1;

Qy          6 FSPSSSSYSSLFT--AKPTRLLSPK 28
              |  |  |||  | ||   :| | ::||
Db        235 FGPEESSYGSPFTPAKRPKRKVAPK 259


RESULT 8
US-08-977-221-34
; Sequence 34, Application US/08977221
; Patent No. 6084069
;  GENERAL INFORMATION:
;    APPLICANT:  UNITED STATES OF AMERICA; DEPT.
;    APPLICANT:  OF HEALTH AND HUMAN SERVICES
;    TITLE OF INVENTION:  MOTILITY STIMULATING
;    TITLE OF INVENTION:  PROTEIN USEFUL IN CANCER DIAGNOSIS AND
;    TITLE OF INVENTION:  THERAPY
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  U.S.A.
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/977,221
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/346,455
;      FILING DATE:  28-NOV-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/249,182
;      FILING DATE:  25-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/822,043
;      FILING DATE:  17-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DOROTHY R. AUTH
;      REGISTRATION NUMBER:  36,434
;      REFERENCE/DOCKET NUMBER:  2026-4149US3
;    TELECOMMUNICATION INFORMATION:
```

```
;       TELEPHONE:  (212) 758-4800
;       TELEFAX:  (212) 751-6849
;   INFORMATION FOR SEQ ID NO:  34:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  829
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  Unknown
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  No
;     ORIGINAL SOURCE:
;       ORGANISM:  Human
;       STRAIN:
;       INDIVIDUAL ISOLATE:
;       DEVELOPMENTAL STAGE:
;       HAPLOTYPE:
;       TISSUE TYPE:
;       CELL TYPE:  Melanoma
;       CELL LINE:  A2058
;       ORGANELLE:
;     FEATURE:
;       NAME/KEY:
;       LOCATION:
;       IDENTIFICATION METHOD:
;       OTHER INFORMATION:  Putative protein
;       OTHER INFORMATION:  sequence of A2058 Autotaxin
US-08-977-221-34

  Query Match             29.9%;  Score 53;  DB 2;  Length 829;
  Best Local Similarity   48.0%;  Pred. No. 59;
  Matches   12;  Conservative    3;  Mismatches    8;  Indels    2;  Gaps    1;

Qy          6 FSPSSSSYSSLFT--AKPTRLLSPK 28
              | |  ||| | ||   :| | ::||
Db        235 FGPEESSYGSPFTPAKRPKRKVAPK 259


RESULT 9
US-09-483-831B-34
; Sequence 34, Application US/09483831B
; Patent No. 6417338
; GENERAL INFORMATION:
;  APPLICANT: STRACKE, MARY
;  APPLICANT:  LIOTTA, LANCE
;  APPLICANT:  SCHIFFMANN, ELLIOTT
;  APPLICANT:  KRUTZCH, HENRY
;  APPLICANT:  MURATA, JUN
;  TITLE OF INVENTION: AUTOTAXIN: MOTILITY STIMULATING PROTEIN USEFUL IN
;  TITLE OF INVENTION:  CANCER DIAGNOSIS AND THERAPY
;  FILE REFERENCE: 2026-4149US4
;  CURRENT APPLICATION NUMBER: US/09/483,831B
;  CURRENT FILING DATE:  2000-01-17
;  PRIOR APPLICATION NUMBER: 07/822,043
;  PRIOR FILING DATE: 1992-01-17
;  PRIOR APPLICATION NUMBER: 08/249,182
;  PRIOR FILING DATE: 1994-05-25
;  PRIOR APPLICATION NUMBER: 08/346,455
;  PRIOR FILING DATE: 1994-11-28
;  PRIOR APPLICATION NUMBER: 08/977,221
;  PRIOR FILING DATE: 1997-11-24
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
;   LENGTH: 829
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Putative protein sequence of A2058 Autotoxin
US-09-483-831B-34

  Query Match             29.9%;  Score 53;  DB 2;  Length 829;
  Best Local Similarity   48.0%;  Pred. No. 59;
  Matches   12;  Conservative    3;  Mismatches    8;  Indels    2;  Gaps    1;

Qy          6 FSPSSSSYSSLFT--AKPTRLLSPK 28
              | |  ||| | ||   :| | ::||
Db        235 FGPEESSYGSPFTPAKRPKRKVAPK 259


RESULT 10
PCT-US95-06613-34
; Sequence 34, Application PC/TUS9506613
;  GENERAL INFORMATION:
;    APPLICANT:  STRACKE, MARY; LIOTTA, LANCE;
;    APPLICANT:  SCHIFFMANN, ELLIOTT; KRUTZSCH,
;    APPLICANT:  HENRY; MURATA, JUN
;    TITLE OF INVENTION:  MOTILITY STIMULATING
;    TITLE OF INVENTION:  PROTEIN USEFUL IN CANCER DIAGNOSIS AND
;    TITLE OF INVENTION:  THERAPY
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  U.S.A.
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/06613
;      FILING DATE:  24-MAY-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/346,455
;      FILING DATE:  28-NOV-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/249,182
;      FILING DATE:  25-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/822,043
;      FILING DATE:  17-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DOROTHY R. AUTH
;      REGISTRATION NUMBER:  36,434
;      REFERENCE/DOCKET NUMBER:  2026-4149US2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  829
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  Unknown
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  No
;    ORIGINAL SOURCE:
;      ORGANISM:  Human
;      STRAIN:
;      INDIVIDUAL ISOLATE:
;      DEVELOPMENTAL STAGE:
;      HAPLOTYPE:
;      TISSUE TYPE:
;      CELL TYPE:  Melanoma
;      CELL LINE:  A2058
;      ORGANELLE:
;    FEATURE:
;      NAME/KEY:
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:  Putative protein
;      OTHER INFORMATION:  sequence of A2058 Autotaxin
```

PCT-US95-06613-34

Query Match 29.9%; Score 53; DB 4; Length 829;
Best Local Similarity 48.0%; Pred. No. 59;
Matches 12; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

```
Qy          6 FSPSSSSYSSLFT--AKPTRLLSPK 28
              | |  ||| | ||   :| | ::||
Db        235 FGPEESSYGSPFTPAKRPKRKVAPK 259
```

RESULT 11
US-08-346-455B-69
; Sequence 69, Application US/08346455B
; Patent No. 5731167
; GENERAL INFORMATION:
; APPLICANT: UNITED STATES OF AMERICA; DEPT.
; APPLICANT: OF HEALTH AND HUMAN SERVICES
; TITLE OF INVENTION: MOTILITY STIMULATING
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND
; TITLE OF INVENTION: THERAPY
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/346,455B
; FILING DATE: 28-NOV-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06613
; FILING DATE: 24-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/249,182
; FILING DATE: 25-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/822,043
; FILING DATE: 17-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36,434
; REFERENCE/DOCKET NUMBER: 2026-4149PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 915
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: Unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: No
; FEATURE:
; NAME/KEY: A2058 ATX protein
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
US-08-346-455B-69

Query Match 29.9%; Score 53; DB 1; Length 915;
Best Local Similarity 48.0%; Pred. No. 67;
Matches 12; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

```
Qy          6 FSPSSSSYSSLFT--AKPTRLLSPK 28
              | |  ||| | ||   :| | ::||
Db        321 FGPEESSYGSPFTPAKRPKRKVAPK 345
```

RESULT 12
US-08-977-221-69
; Sequence 69, Application US/08977221
; Patent No. 6084069
; GENERAL INFORMATION:
; APPLICANT: UNITED STATES OF AMERICA; DEPT.
; APPLICANT: OF HEALTH AND HUMAN SERVICES
; TITLE OF INVENTION: MOTILITY STIMULATING
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND
; TITLE OF INVENTION: THERAPY
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/977,221
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/346,455
; FILING DATE: 28-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/249,182
; FILING DATE: 25-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/822,043
; FILING DATE: 17-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36,434
; REFERENCE/DOCKET NUMBER: 2026-4149US3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 915
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: Unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: No
; FEATURE:
; NAME/KEY: A2058 ATX protein
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
US-08-977-221-69

Query Match 29.9%; Score 53; DB 2; Length 915;
Best Local Similarity 48.0%; Pred. No. 67;
Matches 12; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

```
Qy          6 FSPSSSSYSSLFT--AKPTRLLSPK 28
              | |  ||| | ||   :| | ::||
Db        321 FGPEESSYGSPFTPAKRPKRKVAPK 345
```

RESULT 13

```
US-09-483-831B-69
; Sequence 69, Application US/09483831B
; Patent No. 6417338
; GENERAL INFORMATION:
;  APPLICANT: STRACKE, MARY
;  APPLICANT:  LIOTTA, LANCE
;  APPLICANT:  SCHIFFMANN, ELLIOTT
;  APPLICANT:  KRUTZCH, HENRY
;  APPLICANT:  MURATA, JUN
;  TITLE OF INVENTION: AUTOTAXIN: MOTILITY STIMULATING PROTEIN USEFUL IN
;  TITLE OF INVENTION:  CANCER DIAGNOSIS AND THERAPY
;  FILE REFERENCE: 2026-4149US4
;  CURRENT APPLICATION NUMBER: US/09/483,831B
;  CURRENT FILING DATE:  2000-01-17
;  PRIOR APPLICATION NUMBER: 07/822,043
;  PRIOR FILING DATE: 1992-01-17
;  PRIOR APPLICATION NUMBER: 08/249,182
;  PRIOR FILING DATE: 1994-05-25
;  PRIOR APPLICATION NUMBER: 08/346,455
;  PRIOR FILING DATE: 1994-11-28
;  PRIOR APPLICATION NUMBER: 08/977,221
;  PRIOR FILING DATE: 1997-11-24
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 69
;   LENGTH: 915
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
;   OTHER INFORMATION: Polypeptide
US-09-483-831B-69

  Query Match             29.9%;  Score 53;  DB 2;  Length 915;
  Best Local Similarity   48.0%;  Pred. No. 67;
  Matches   12;  Conservative    3;  Mismatches    8;  Indels    2;  Gaps    1;

Qy          6 FSPSSSSYSSLFT--AKPTRLLSPK 28
              | |  ||| | ||   :| | ::||
Db        321 FGPEESSYGSPFTPAKRPKRKVAPK 345


RESULT 14
PCT-US95-06613-69
; Sequence 69, Application PC/TUS9506613
;  GENERAL INFORMATION:
;    APPLICANT:  STRACKE, MARY; LIOTTA, LANCE;
;    APPLICANT:  SCHIFFMANN, ELLIOTT; KRUTZSCH,
;    APPLICANT:  HENRY; MURATA, JUN
;    TITLE OF INVENTION:  MOTILITY STIMULATING
;    TITLE OF INVENTION:  PROTEIN USEFUL IN CANCER DIAGNOSIS AND
;    TITLE OF INVENTION:  THERAPY
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  U.S.A.
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/06613
;      FILING DATE:  24-MAY-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/346,455
;      FILING DATE:  28-NOV-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/249,182
;      FILING DATE:  25-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/822,043
;      FILING DATE:  17-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DOROTHY R. AUTH
;      REGISTRATION NUMBER:  36,434
;      REFERENCE/DOCKET NUMBER:  2026-4149US2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;  INFORMATION FOR SEQ ID NO:  69:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  915
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  Unknown
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  No
;    FEATURE:
;      NAME/KEY:  A2058 ATX protein
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
PCT-US95-06613-69

  Query Match             29.9%;  Score 53;  DB 4;  Length 915;
  Best Local Similarity   48.0%;  Pred. No. 67;
  Matches   12;  Conservative    3;  Mismatches    8;  Indels    2;  Gaps    1;

Qy          6 FSPSSSSYSSLFT--AKPTRLLSPK 28
              | |  ||| | ||   :| | ::||
Db        321 FGPEESSYGSPFTPAKRPKRKVAPK 345


RESULT 15
US-09-914-272A-1
; Sequence 1, Application US/09914272A
; Patent No. 6673913
; GENERAL INFORMATION:
;  APPLICANT: Sakaguchi, No. 6673913uo
;  APPLICANT:  Kuwahara, Kazuhiko
;  TITLE OF INVENTION: GANP Protein
;  FILE REFERENCE: 050208-0014
;  CURRENT APPLICATION NUMBER: US/09/914,272A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: PCT/JP99/04634
;  PRIOR FILING DATE: 1999-08-27
;  PRIOR APPLICATION NUMBER: 47035/1999
;  PRIOR FILING DATE: 1999-02-24
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1971
;   TYPE: PRT
;   ORGANISM: Mouse
US-09-914-272A-1

  Query Match             29.9%;  Score 53;  DB 2;  Length 1971;
  Best Local Similarity   38.9%;  Pred. No. 1.7e+02;
  Matches   14;  Conservative    6;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36
              : :| |  |:||:|:  || |  |  | |    |:|
Db        231 VSTSAFGSSNSSFSTFPTASPGSLGEPFPANKPSLR 266


Search completed: March  7, 2006, 22:00:14
Job time : 40.2 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2006, 22:00:32 ; Search time 15.6 Seconds
(without alignments)
46.155 Million cell updates/sec

Title: US-10-751-235-11
Perfect score: 177
Sequence: 1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 135346 seqs, 20000420 residues

Total number of hits satisfying chosen parameters: 135346

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	53	29.9	352	6	US-10-838-616-4	Sequence 4, Appli
2	53	29.9	639	7	US-11-094-586-2	Sequence 2, Appli
3	53	29.9	1971	7	US-11-179-624-1	Sequence 1, Appli
4	48.5	27.4	505	7	US-11-087-099-6346	Sequence 6346, Ap
5	48.5	27.4	1905	6	US-10-877-346-44	Sequence 44, Appl
6	48	27.1	527	7	US-11-240-341-44	Sequence 44, Appl
7	48	27.1	1053	6	US-10-517-939-52	Sequence 52, Appl
8	48	27.1	1142	7	US-11-044-051-73	Sequence 73, Appl
9	47.5	26.8	312	7	US-11-072-512-3588	Sequence 3588, Ap
10	47.5	26.8	797	6	US-10-514-531-5	Sequence 5, Appli
11	47.5	26.8	798	6	US-10-514-531-2	Sequence 2, Appli
12	47	26.6	130	6	US-10-793-626-164	Sequence 164, App
13	47	26.6	161	7	US-11-087-099-2465	Sequence 2465, Ap
14	47	26.6	688	7	US-11-165-226-124	Sequence 124, App
15	47	26.6	906	7	US-11-087-099-11997	Sequence 11997, A
16	47	26.6	1199	6	US-10-821-234-1126	Sequence 1126, Ap
17	47	26.6	1972	7	US-11-124-367A-446	Sequence 446, App
18	46.5	26.3	199	6	US-10-714-887-238	Sequence 238, App
19	46.5	26.3	285	7	US-11-072-512-3028	Sequence 3028, Ap
20	46.5	26.3	680	6	US-10-467-962B-101	Sequence 101, App
21	46	26.0	1890	7	US-11-033-039-314	Sequence 314, App
22	45	25.4	120	6	US-10-915-161-20	Sequence 20, Appl
23	45	25.4	366	6	US-10-524-647-126	Sequence 126, App
24	45	25.4	366	6	US-10-524-972-114	Sequence 114, App
25	45	25.4	565	6	US-10-915-161-2	Sequence 2, Appli
26	45	25.4	723	7	US-11-072-512-2246	Sequence 2246, Ap
27	45	25.4	2725	7	US-11-113-424-52	Sequence 52, Appl
28	45	25.4	2725	7	US-11-100-640-10	Sequence 10, Appl
29	45	25.4	2725	7	US-11-100-640-16	Sequence 16, Appl
30	44.5	25.1	175	6	US-10-821-234-1543	Sequence 1543, Ap
31	44.5	25.1	179	7	US-11-197-133A-62	Sequence 62, Appl
32	44.5	25.1	838	7	US-11-114-906-40	Sequence 40, Appl
33	44.5	25.1	851	7	US-11-114-906-38	Sequence 38, Appl
34	44.5	25.1	863	7	US-11-114-906-32	Sequence 32, Appl
35	44.5	25.1	876	7	US-11-114-906-30	Sequence 30, Appl
36	44.5	25.1	951	7	US-11-114-906-36	Sequence 36, Appl
37	44.5	25.1	957	7	US-11-114-906-34	Sequence 34, Appl
38	44.5	25.1	976	7	US-11-114-906-28	Sequence 28, Appl
39	44.5	25.1	982	7	US-11-114-906-26	Sequence 26, Appl
40	44.5	25.1	1027	7	US-11-024-959-324	Sequence 324, App
41	44.5	25.1	1730	7	US-11-182-016-19	Sequence 19, Appl
42	44	24.9	384	7	US-11-134-563-2	Sequence 2, Appli
43	44	24.9	2344	6	US-10-330-773-627	Sequence 627, App
44	43.5	24.6	71	7	US-11-207-078-331	Sequence 331, App
45	43.5	24.6	108	6	US-10-793-626-2898	Sequence 2898, Ap

ALIGNMENTS

RESULT 1
US-10-838-616-4
; Sequence 4, Application US/10838616
; Publication No. US20060008874A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: CREELMAN, Robert A
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: KUMIMOTO, Roderick W
; APPLICANT: GUTTERSON, Neal I
; APPLICANT: REUBER, T. Lynne
; APPLICANT: LIBBY, Jeffrey M
; TITLE OF INVENTION: Plant Transcriptional Regulators of Abiotic Stress
; FILE REFERENCE: MBI-0069CIP
; CURRENT APPLICATION NUMBER: US/10/838,616
; CURRENT FILING DATE: 2004-05-04
; PRIOR APPLICATION NUMBER: Stress-Related Polypeptides in Plants
; PRIOR FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 10/685,922
; PRIOR FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 09/810,836
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 10/412,699
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G9 polypeptide Paralogous to G867, G993, G1930
US-10-838-616-4

Query Match 29.9%; Score 53; DB 6; Length 352;

Best Local Similarity 44.8%; Pred. No. 5.2;
Matches 13; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

```
Qy          1 MESSLFSPSSSSYSSLFTAKPTRLLSPKP 29
              |:||     ||| |  |:|   : ||| |
Db          1 MDSSCIDEISSSTSESFSATTAKKLSPPP 29
```

RESULT 2
US-11-094-586-2
; Sequence 2, Application US/11094586
; Publication No. US20050273886A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Butler, Karlene H.
; APPLICANT: Carlson, Thomas J.
; APPLICANT: Hitz, William D.
; APPLICANT: Stoop, Johan M.
; TITLE OF INVENTION: Plastidic Phosphoglucomutase Genes
; FILE REFERENCE: BB1451 USCIP
; CURRENT APPLICATION NUMBER: US/11/094,586
; CURRENT FILING DATE: 2005-03-30
; PRIOR APPLICATION NUMBER: US 09/906,209
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: US 60/218,712
; PRIOR FILING DATE: 2000-07-17
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Microsoft Office 97 & PatentIn Version 3.3
; SEQ ID NO 2
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Typha latifolia
US-11-094-586-2

Query Match 29.9%; Score 53; DB 7; Length 639;
Best Local Similarity 38.2%; Pred. No. 10;
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

```
Qy          3 SSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36
              | | |||||| : | :::   |   :|   ||::
Db         37 SKLLSPSSSSPAVLVSSRIPLLSLRRPNLRFSVK 70
```

RESULT 3
US-11-179-624-1
; Sequence 1, Application US/11179624
; Publication No. US20060040372A1
; GENERAL INFORMATION:
; APPLICANT: Sakaguchi, Nobuo
; APPLICANT: Kuwahara, Kazuhiko
; TITLE OF INVENTION: GANP Protein
; FILE REFERENCE: 050208-0018
; CURRENT APPLICATION NUMBER: US/11/179,624
; CURRENT FILING DATE: 2005-07-13
; PRIOR APPLICATION NUMBER: PCT/JP99/04634
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 09/914,272
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 10/747,133
; PRIOR FILING DATE: 2003-12-30
; PRIOR APPLICATION NUMBER: 47035/1999
; PRIOR FILING DATE: 1999-02-24
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 1971
; TYPE: PRT
; ORGANISM: Mouse
US-11-179-624-1

Query Match 29.9%; Score 53; DB 7; Length 1971;
Best Local Similarity 38.9%; Pred. No. 35;
Matches 14; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

```
Qy          1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36
              : :| |  |:||:|:  || |  |  | |    |:|
Db        231 VSTSAFGSSNSSFSTFPTASPGSLGEPFPANKPSLR 266
```

RESULT 4
US-11-087-099-6346
; Sequence 6346, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 6346
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Neurospora crassa
US-11-087-099-6346

Query Match 27.4%; Score 48.5; DB 7; Length 505;
Best Local Similarity 40.0%; Pred. No. 32;
Matches 14; Conservative 4; Mismatches 14; Indels 3; Gaps 1;

```
Qy          2 ESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36
              :||||:| :  ||     :  ||  | |   | ||
Db        333 DSSLFNPVTMPYSMAVFYEVLRLYPPIP---FEIR 364
```

RESULT 5
US-10-877-346-44
; Sequence 44, Application US/10877346
; Publication No. US20060014153A1
; GENERAL INFORMATION:
; APPLICANT: Gerlach, Valerie L
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Leach, Martin D
; APPLICANT: Shimkets, Richard A
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-124
; CURRENT APPLICATION NUMBER: US/10/877,346
; CURRENT FILING DATE: 2004-06-25
; PRIOR APPLICATION NUMBER: US/09/964,956
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/235,631
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/235,633
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/235,808
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,064
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,065
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,066
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,135
; PRIOR FILING DATE: 2000-09-28

```
;  PRIOR APPLICATION NUMBER: 60/237,434
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/238,321
;  PRIOR FILING DATE: 2000-10-05
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 127
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
;    LENGTH: 1905
;    TYPE: PRT
;    ORGANISM: Xenopus laevis
US-10-877-346-44

  Query Match             27.4%;  Score 48.5;  DB 6;  Length 1905;
  Best Local Similarity   38.7%;  Pred. No. 1.4e+02;
  Matches   12;  Conservative    5;  Mismatches   13;  Indels    1;  Gaps    1;

Qy          1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKF 31
              ::: | || |:     ||:|  ||    |||
Db        254 LDTQLTSPDSTG-EQFFTSKIVRLCVDDPKF 283


RESULT 6
US-11-240-341-44
; Sequence 44, Application US/11240341
; Publication No. US20060024742A1
; GENERAL INFORMATION:
;  APPLICANT: Martelange, Valerie
;  APPLICANT:  De Smet, Charles
;  APPLICANT:  Boon-Falleur, Thierry
;  TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR
;  FILE REFERENCE: L0461.70047US01
;  CURRENT APPLICATION NUMBER: US/11/240,341
;  CURRENT FILING DATE:  2005-09-30
;  PRIOR APPLICATION NUMBER: US 09/183,789
;  PRIOR FILING DATE: 1998-10-30
;  PRIOR APPLICATION NUMBER: US 09/060,706
;  PRIOR FILING DATE: 1998-04-15
;  NUMBER OF SEQ ID NOS: 47
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 44
;    LENGTH: 527
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-11-240-341-44

  Query Match             27.1%;  Score 48;  DB 7;  Length 527;
  Best Local Similarity   42.9%;  Pred. No. 39;
  Matches   12;  Conservative    5;  Mismatches    7;  Indels    4;  Gaps    1;

Qy          8 PSSSSYSSLFTAKPT----RLLSPKPKF 31
              |:: | ::||  |||     |:  ||||
Db        496 PTTESEAALFGDKPTIKQPMLILRKPKF 523


RESULT 7
US-10-517-939-52
; Sequence 52, Application US/10517939
; Publication No. US20060003433A1
; GENERAL INFORMATION:
;  APPLICANT: Steer, Brian
;  APPLICANT:  Callen, Walter
;  APPLICANT:  Healey, Shaun
;  APPLICANT:  Hazlewood, Geoff
;  APPLICANT:  Wu, Di
;  APPLICANT:  Blum, David
;  APPLICANT:  Esteghlalian, Alireza
;  TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM
;  TITLE OF INVENTION:  AND METHODS FOR MAKING AND USING THEM
;  FILE REFERENCE: 564462007901
;  CURRENT APPLICATION NUMBER: US/10/517,939
;  CURRENT FILING DATE:  2004-12-13
;  PRIOR APPLICATION NUMBER: PCT/US03/19153
;  PRIOR FILING DATE: 2003-06-16
;  PRIOR APPLICATION NUMBER: 60/389,299
;  PRIOR FILING DATE: 2002-06-14
;  NUMBER OF SEQ ID NOS: 380
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 52
;    LENGTH: 1053
;    TYPE: PRT
;    ORGANISM: Unknown
;    FEATURE:
;    OTHER INFORMATION: Obtained from an environmental sample
;    FEATURE:
;    NAME/KEY: SIGNAL
;    LOCATION: (1)...(30)
US-10-517-939-52

  Query Match             27.1%;  Score 48;  DB 6;  Length 1053;
  Best Local Similarity   42.9%;  Pred. No. 84;
  Matches   15;  Conservative    4;  Mismatches    4;  Indels   12;  Gaps    3;

Qy          6 FSPSSSSYSSLFTAKPTR----LLSPK-PKFTFSI 35
              ||||       | : ||:    ::||| | ||| |
Db        159 FSPS-------FESTPTKCSLIVVSPKNPSFTFYI 186


RESULT 8
US-11-044-051-73
; Sequence 73, Application US/11044051
; Publication No. US20050255553A1
; GENERAL INFORMATION:
;  APPLICANT: VAN PEL, Aline
;  APPLICANT:  GODELAINE, DaniSle
;  APPLICANT:  CARRASCO, Javier
;  APPLICANT:  BRASSEUR, Francis
;  APPLICANT:  BOON-FALLEUR, Thierry
;  TITLE OF INVENTION: MAGE C2 ANTIGENIC PEPTIDES AND USES THEREOF
;  FILE REFERENCE: LUD 5888 US
;  CURRENT APPLICATION NUMBER: US/11/044,051
;  CURRENT FILING DATE:  2005-01-28
;  PRIOR APPLICATION NUMBER: US 60/459,263
;  PRIOR FILING DATE: 2004-02-09
;  NUMBER OF SEQ ID NOS: 111
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 73
;    LENGTH: 1142
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-11-044-051-73

  Query Match             27.1%;  Score 48;  DB 7;  Length 1142;
  Best Local Similarity   33.3%;  Pred. No. 92;
  Matches   13;  Conservative    9;  Mismatches   13;  Indels    4;  Gaps    1;

Qy          1 MESSLFSPSSSSYS----SLFTAKPTRLLSPKPKFTFSI 35
              ::| | :|:|| :|    |:| : |  : ||   |  |:
Db        124 VQSPLQNPASSFFSSALLSIFQSSPESIQSPFEGFPQSV 162


RESULT 9
US-11-072-512-3588
; Sequence 3588, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
```

```
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: 084335-0191
;  CURRENT APPLICATION NUMBER: US/11/072,512
;  CURRENT FILING DATE:  2005-03-07
;  PRIOR APPLICATION NUMBER: US 60/350,978
;  PRIOR FILING DATE: 2002-01-25
;  PRIOR APPLICATION NUMBER: JP 2001-379298
;  PRIOR FILING DATE: 2001-11-05
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3588
;   LENGTH: 312
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-072-512-3588

  Query Match             26.8%;  Score 47.5;  DB 7;  Length 312;
  Best Local Similarity   40.7%;  Pred. No. 25;
  Matches   11;  Conservative    4;  Mismatches    9;  Indels    3;  Gaps    1;

Qy          7 SPSSSSYSSLFTA---KPTRLLSPKPK 30
              || |   ||:|::    ||   || |:
Db        238 SPDSEGLSSVFSSSLPSPTNSSSPSPR 264


RESULT 10
US-10-514-531-5
; Sequence 5, Application US/10514531
; Publication No. US20060035849A1
; GENERAL INFORMATION:
;  APPLICANT: Spiegelman, Bruce M., et al.
;  TITLE OF INVENTION: Methods and Compositions For Modulating Type I Muscle
;  TITLE OF INVENTION:  Formation Using PGC-1alpha
;  FILE REFERENCE: DFN-041US
;  CURRENT APPLICATION NUMBER: US/10/514,531
;  CURRENT FILING DATE:  2004-11-12
;  PRIOR APPLICATION NUMBER: PCT/US03/04792
;  PRIOR FILING DATE: 2003-02-13
;  PRIOR APPLICATION NUMBER: 60/357069
;  PRIOR FILING DATE: 2002-02-13
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 797
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-514-531-5

  Query Match             26.8%;  Score 47.5;  DB 6;  Length 797;
  Best Local Similarity   35.7%;  Pred. No. 72;
  Matches   15;  Conservative    6;  Mismatches    8;  Indels   13;  Gaps    2;

Qy          4 SLF--SPSSSSYS-----------SLFTAKPTRLLSPKPKFT 32
              |||  ||| ||::           |||: :| |: |    |:
Db        531 SLFDVSPSCSSFNSPCRDSVSPPKSLFSQRPQRMRSRSRSFS 572


RESULT 11
US-10-514-531-2
; Sequence 2, Application US/10514531
; Publication No. US20060035849A1
; GENERAL INFORMATION:
;  APPLICANT: Spiegelman, Bruce M., et al.
;  TITLE OF INVENTION: Methods and Compositions For Modulating Type I Muscle
;  TITLE OF INVENTION:  Formation Using PGC-1alpha
;  FILE REFERENCE: DFN-041US
;  CURRENT APPLICATION NUMBER: US/10/514,531
;  CURRENT FILING DATE:  2004-11-12
;  PRIOR APPLICATION NUMBER: PCT/US03/04792
;  PRIOR FILING DATE: 2003-02-13
;  PRIOR APPLICATION NUMBER: 60/357069
;  PRIOR FILING DATE: 2002-02-13
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 798
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-514-531-2

  Query Match             26.8%;  Score 47.5;  DB 6;  Length 798;
  Best Local Similarity   35.7%;  Pred. No. 72;
  Matches   15;  Conservative    6;  Mismatches    8;  Indels   13;  Gaps    2;

Qy          4 SLF--SPSSSSYS-----------SLFTAKPTRLLSPKPKFT 32
              |||  ||| ||::           |||: :| |: |    |:
Db        532 SLFNVSPSCSSFNSPCRDSVSPPKSLFSQRPQRMRSRSRSFS 573


RESULT 12
US-10-793-626-164
; Sequence 164, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
;  APPLICANT: KIMMERLY, WILLIAM JOHN
;  TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
;  FILE REFERENCE: PU3480US
;  CURRENT APPLICATION NUMBER: US/10/793,626
;  CURRENT FILING DATE:  2004-03-04
;  PRIOR APPLICATION NUMBER: 60/164,258
;  PRIOR FILING DATE: 1999-11-09
;  NUMBER OF SEQ ID NOS: 4472
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 164
;   LENGTH: 130
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: amino acid sequence
US-10-793-626-164

  Query Match             26.6%;  Score 47;  DB 6;  Length 130;
  Best Local Similarity   42.4%;  Pred. No. 11;
  Matches   14;  Conservative    5;  Mismatches    4;  Indels   10;  Gaps    2;

Qy          3 SSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSI 35
              |::  |||| |:|:|          ||| | ||
Db         66 STVHKPSSST-STLYT---------KPKLTISI 88


RESULT 13
US-11-087-099-2465
; Sequence 2465, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
;  APPLICANT: Abad, Mark S. et al.
;  TITLE OF INVENTION: Genes and Uses for Plant Improvement
;  FILE REFERENCE: 38-21(53450)B EP
;  CURRENT APPLICATION NUMBER: US/11/087,099
;  CURRENT FILING DATE:  2005-03-22
;  NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 2465
;   LENGTH: 161
;   TYPE: PRT
;   ORGANISM: Oryza sativa (japonica cultivar-group)
US-11-087-099-2465
```

```
  Query Match            26.6%;  Score 47;  DB 7;  Length 161;
  Best Local Similarity  31.7%;  Pred. No. 14;
  Matches   13;  Conservative    6;  Mismatches   10;  Indels   12;  Gaps    1;

Qy          1 MESSLFSPSSSSYSSLFTAKPTRLLS-----------PKP 29
              :|   |:|:||| : |  |||  : :            |||
Db         57 VEKKAFTPASSSKTPLEPAKPDEVPAAAVEDKQAAQEPPKP 97


RESULT 14
US-11-165-226-124
; Sequence 124, Application US/11165226
; Publication No. US20050287627A1
; GENERAL INFORMATION:
;  APPLICANT: KYOWA HAKKO KOGYO CO., LTD.
;  TITLE OF INVENTION: Process for producing dipeptides or dipeptide derivatives
;  FILE REFERENCE: 4093-13
;  CURRENT APPLICATION NUMBER: US/11/165,226
;  CURRENT FILING DATE:  2005-06-24
;  PRIOR APPLICATION NUMBER: JP2004-189007
;  PRIOR FILING DATE: 2004-06-25
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 124
;   LENGTH: 688
;   TYPE: PRT
;   ORGANISM: Eschierichia coli W3110
US-11-165-226-124

  Query Match            26.6%;  Score 47;  DB 7;  Length 688;
  Best Local Similarity  44.0%;  Pred. No. 71;
  Matches   11;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 MESSLFSPSSSSYSSLFTAKPTRLL 25
              ||:||    |||     ||:| | :
Db        239 MEASLMELMSSSLKQRLTAEPVRFV 263


RESULT 15
US-11-087-099-11997
; Sequence 11997, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
;  APPLICANT: Abad, Mark S. et al.
;  TITLE OF INVENTION: Genes and Uses for Plant Improvement
;  FILE REFERENCE: 38-21(53450)B EP
;  CURRENT APPLICATION NUMBER: US/11/087,099
;  CURRENT FILING DATE:  2005-03-22
;  NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 11997
;   LENGTH: 906
;   TYPE: PRT
;   ORGANISM: Pseudomonas syringae pv. syringae B728a
US-11-087-099-11997

  Query Match            26.6%;  Score 47;  DB 7;  Length 906;
  Best Local Similarity  39.4%;  Pred. No. 97;
  Matches   13;  Conservative    7;  Mismatches    7;  Indels    6;  Gaps    3;

Qy          6 FSPSSSSYSSLFTA--KPTRLLSPKPKFTFSIR 36
              :|||: |:|: ||   ||| :  |:    | :|
Db        490 YSPSAGSFSARFTGKIKPT-ITGPQ---VFKVR 518


Search completed: March  7, 2006, 22:06:39
Job time : 16.6 secs
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2006, 21:53:39 ; Search time 28.8 Seconds
(without alignments)
120.271 Million cell updates/sec

Title: US-10-751-235-11
Perfect score: 177
Sequence: 1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	177	100.0	566	2	T46159	cytochrome P450-li
2	65.5	37.0	430	2	T50296	conserved hypothet
3	60	33.9	355	1	JQ2256	phosphoribosylform
4	60	33.9	389	2	T47640	phosphoribosylform
5	59.5	33.6	319	2	A85437	hypothetical prote
6	58.5	33.1	351	2	T19059	hypothetical prote
7	56	31.6	474	2	T45818	hypothetical prote
8	56	31.6	883	2	S04722	puff 74E protein -
9	55	31.1	406	2	T48103	mRNA binding prote
10	54.5	30.8	612	2	T38714	hypothetical prote
11	54.5	30.8	1621	2	T30200	protein-tyrosine k
12	54	30.5	132	2	A75491	hypothetical prote
13	54	30.5	205	2	T47591	aintegumaenta-like
14	54	30.5	403	2	T45712	NAD-dependent mala
15	54	30.5	403	2	T51862	malate dehydrogena
16	53.5	30.2	1564	2	S55517	probable transport
17	53	29.9	352	2	T51330	DNA binding protei
18	53	29.9	915	1	A55144	autotaxin precurso
19	52.5	29.7	402	2	A84766	hypothetical prote
20	52.5	29.7	411	2	T51818	ATP phosphoribosyl
21	52	29.4	159	2	D71033	hypothetical prote
22	52	29.4	246	2	T30490	hypothetical prote
23	52	29.4	320	2	T31547	hypothetical prote
24	52	29.4	2606	2	T03159	large tegument pro
25	51.5	29.1	272	2	H71124	hypothetical prote
26	51	28.8	211	2	C84751	hypothetical prote
27	51	28.8	299	2	T52452	ATP-dependent Clp
28	51	28.8	354	2	T49806	hypothetical prote
29	51	28.8	446	2	T45525	WSC4 homolog [impo
30	51	28.8	499	2	AH3323	UDP-N-acetylmuramo
31	51	28.8	620	2	S52494	protein kinase hom
32	51	28.8	2305	2	B89608	protein C23F12.1 [
33	51	28.8	2305	2	T15571	hypothetical prote
34	50.5	28.5	361	2	T01998	hypothetical prote
35	50.5	28.5	632	2	T00679	hypothetical prote
36	50.5	28.5	960	2	S54461	hypothetical prote
37	50.5	28.5	1390	2	T18883	hypothetical prote
38	50	28.2	291	2	E84849	hypothetical prote
39	50	28.2	562	2	T05758	hypothetical prote
40	50	28.2	687	2	AE0345	polyphosphate kina
41	50	28.2	688	2	AC0819	polyphosphate kina
42	50	28.2	707	2	T14195	extensin homolog T
43	50	28.2	848	2	E86443	probable G-protein
44	50	28.2	1695	2	T19823	hypothetical prote
45	49.5	28.0	113	2	F72746	hypothetical prote

ALIGNMENTS

RESULT 1
T46159
cytochrome P450-like protein - Arabidopsis thaliana
N;Alternate names: protein T4D2.60
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C;Accession: T46159
R;Nyakatura, G.; Fartmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichselgartner, M.; M
submitted to the Protein Sequence Database, December 1999
A;Reference number: Z23025
A;Accession: T46159
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-566 <NYA>
A;Cross-references: UNIPROT:Q9SCP8; UNIPARC:UPI000009D7F9; EMBL:AL132958
A;Experimental source: cultivar Columbia; BAC clone T4D2
C;Genetics:
A;Map position: 3
A;Introns: 183/2; 292/3; 358/3; 392/3; 439/3; 475/3; 503/2; 557/3
A;Note: T4D2.60
C;Superfamily: pea cytochrome P450 CYP97; cytochrome P450 homology
C;Keywords: heme; iron; metalloprotein
F;371-536/Domain: cytochrome P450 homology <P45>
F;514/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 100.0%; Score 177; DB 2; Length 566;
Best Local Similarity 100.0%; Pred. No. 1e-15;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36
||||||||||||||||||||||||||||||||||||
Db 1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36

RESULT 2
T50296
conserved hypothetical SH3 domain protein [imported] - fission yeast (Schizosaccharomyces
C;Species: Schizosaccharomyces pombe
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jul-2004
C;Accession: T50296
R;McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Brown, S.; Harris, D.
submitted to the EMBL Data Library, December 1999
A;Reference number: Z25040
A;Accession: T50296
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-430 <MCD>
A;Cross-references: UNIPROT:Q9URW6; UNIPARC:UPI000006AE4A; EMBL:AL133359; PIDN:CAB62422.1
A;Experimental source: strain 972h(-)
C;Genetics:
A;Gene: SPDB:SPAPJ696.02
A;Map position: 1

```
A;Introns: 13/1; 26/3

  Query Match             37.0%;  Score 65.5;  DB 2;  Length 430;
  Best Local Similarity   55.2%;  Pred. No. 0.51;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    5;  Gaps    1;

Qy          3 SSLFSPSSSSYSSLFTAKPTRLLSPKPKF 31
              || ||  || ||     ||:|  :|||||
Db        343 SSQFSSRSSEYS-----KPSRPTAPKPKF 366


RESULT 3
JQ2256
phosphoribosylformylglycinamidine cyclo-ligase (EC 6.3.3.1) precursor - Arabidopsis thal
N;Alternate names: 5'-phosphoribosyl-5-aminoimidazole synthetase; protein F28P10.10
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Sep-1993 #sequence_revision 28-Oct-1994 #text_change 15-Oct-1999
C;Accession: JQ2256; T06719
R;Senecoff, J.F.; Meagher, R.B.
Plant Physiol. 102, 387-399, 1993
A;Title: Isolating the Arabidopsis thaliana genes for de novo purine synthesis by suppre
A;Reference number: JQ2256; MUID:94151435; PMID:8108507
A;Accession: JQ2256
A;Molecule type: DNA
A;Residues: 1-355 <SEN>
A;Cross-references: UNIPARC:UPI0000048D0B; GB:L12457; NID:g289187; PIDN:AAC37341.1; PID:
R;Quetier, F.; Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artigu
submitted to the Protein Sequence Database, April 1999
A;Reference number: Z15793
A;Accession: T06719
A;Molecule type: DNA
A;Residues: 1-125 <QUE>
A;Cross-references: UNIPARC:UPI000016DB1E; EMBL:AL049655; ATSP:F28P10.10; GSPDB:GN00061
A;Experimental source: cultivar Columbia; BAC clone F28P10
C;Comment: This enzyme is involved in the fifth step of the purine biosynthetic pathway,
C;Genetics:
A;Gene: ATSP:F28P10.10
A;Map position: 3
A;Introns: 102/1
C;Superfamily: phosphoribosylformylglycinamidine cyclo-ligase; phosphoribosylformylglyci
C;Keywords: cyclo-ligase; purine nucleotide biosynthesis
F;1-58/Domain: signal sequence #status predicted <SIG>
F;59-355/Product: phosphoribosylformylglycinamidine cyclo-ligase #status predicted <MAT>
F;68-355/Domain: phosphoribosylformylglycinamidine cyclo-ligase homology <PFCL>

  Query Match             33.9%;  Score 60;  DB 1;  Length 355;
  Best Local Similarity   42.1%;  Pred. No. 2.2;
  Matches   16;  Conservative    6;  Mismatches   10;  Indels    6;  Gaps    2;

Qy          1 MESSLFSPSSSSYSSLFTAKPTR---LLSPKPKFTFSI 35
              ||: :   ||| ||||:    :|   : ||||   ||:
Db          1 MEARILQSSSSCYSSLYAVNRSRFSSVSSPKP---FSV 35


RESULT 4
T47640
phosphoribosylformylglycinamidine cyclo-ligase precursor - Arabidopsis thaliana
N;Alternate names: protein T15C9.10
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T47640
R;Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24470
A;Accession: T47640
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-389 <MEW>
A;Cross-references: UNIPROT:Q05728; UNIPARC:UPI00000017F8; EMBL:AL132970
A;Experimental source: cultivar Columbia; BAC clone T15C9
C;Genetics:
A;Map position: 3
A;Introns: 102/1; 133/3; 167/3; 187/3; 245/2; 279/3; 332/3
A;Note: T15C9.10
C;Superfamily: phosphoribosylformylglycinamidine cyclo-ligase; phosphoribosylformylglycir

  Query Match             33.9%;  Score 60;  DB 2;  Length 389;
  Best Local Similarity   42.1%;  Pred. No. 2.5;
  Matches   16;  Conservative    6;  Mismatches   10;  Indels    6;  Gaps    2;

Qy          1 MESSLFSPSSSSYSSLFTAKPTR---LLSPKPKFTFSI 35
              ||: :   ||| ||||:    :|   : ||||   ||:
Db          1 MEARILQSSSSCYSSLYAVNRSRFSSVSSPKP---FSV 35


RESULT 5
A85437
hypothetical protein AT4g37000 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C;Accession: A85437
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: A85437
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-319 <STO>
A;Cross-references: UNIPARC:UPI0000048D08; GB:NC_001268; NID:g7270649; PIDN:CAB80366.1; C
C;Genetics:
A;Gene: AT4g37000
A;Map position: 4

  Query Match             33.6%;  Score 59.5;  DB 2;  Length 319;
  Best Local Similarity   43.8%;  Pred. No. 2.3;
  Matches   14;  Conservative    7;  Mismatches    8;  Indels    3;  Gaps    1;

Qy          3 SSLFSPSSSSYSSLFTAKPTRL---LSPKPKF 31
              ::|:| || || |  |:||:|    | |: :|
Db          7 NTLYSSSSPSYLSPLTSKPSRFSKNLRPRAQF 38


RESULT 6
T19059
hypothetical protein T09F3.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T19059; T24756
R;Matthews, P.
submitted to the EMBL Data Library, June 1995
A;Reference number: Z19066
A;Accession: T19059
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-351 <WIL>
A;Cross-references: UNIPROT:Q17793; UNIPARC:UPI000007AE93; EMBL:Z49908; PIDN:CAA90104.1;
A;Experimental source: clone C07E3
R;Lloyd, C.
submitted to the EMBL Data Library, April 1995
A;Reference number: Z19932
A;Accession: T24756
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-351 <WI2>
A;Cross-references: UNIPARC:UPI000007AE93; EMBL:Z49070; PIDN:CAA88872.1; GSPDB:GN00020; C
A;Experimental source: clone T09F3
C;Genetics:
A;Gene: CESP:T09F3.1
A;Map position: 2
A;Introns: 20/3; 122/3; 188/3; 241/2; 274/2

  Query Match             33.1%;  Score 58.5;  DB 2;  Length 351;
  Best Local Similarity   46.2%;  Pred. No. 3.5;
  Matches   18;  Conservative    5;  Mismatches    9;  Indels    7;  Gaps    2;
```

```
Qy          3 SSLFSPSSSS----YSSLFTAK---PTRLLSPKPKFTFS 34
              ||:|: || |    ||:: ||    || | ||| :|  |
Db         73 SSMFAASSPSAAATYSTVTTAALVVPTTLQSPKREFVCS 111
```

RESULT 7
T45818
hypothetical protein F28O9.250 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C;Accession: T45818
R;Benes, V.; Rechmann, S.; Borkova, D.; Ansorge, W.; Mewes, H.W.; Lemcke, K.; Mayer, K.F
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23014
A;Accession: T45818
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-474 <BEN>
A;Cross-references: UNIPROT:Q9M2K7; UNIPARC:UPI00000A7FBF; EMBL:AL137080
A;Experimental source: cultivar Columbia; BAC clone F28O9
C;Genetics:
A;Map position: 3
A;Note: F28O9.250
C;Superfamily: Arabidopsis thaliana hypothetical protein F28O9.250

```
  Query Match             31.6%;  Score 56;  DB 2;  Length 474;
  Best Local Similarity   48.5%;  Pred. No. 10;
  Matches   16;  Conservative    5;  Mismatches    8;  Indels    4;  Gaps    1;

Qy          7 SPSSSSYS----SLFTAKPTRLLSPKPKFTFSI 35
              ||||||  |    :|| :: :|||  | : | ||
Db         33 SPSSSSSSPATTTLFRSRSSRLLLSKAQSTISI 65
```

RESULT 8
S04722
puff 74E protein - fruit fly (Drosophila melanogaster)
N;Alternate names: ecdysone-induced protein E74B; ets-related protein E74B
C;Species: Drosophila melanogaster
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C;Accession: S04722; B34692
R;Janknecht, R.; Taube, W.; Luedecke, H.J.; Pongs, O.
Nucleic Acids Res. 17, 4455-4464, 1989
A;Title: Characterization of a putative transcription factor gene expressed in the 20-OH
A;Reference number: S04722; MUID:89315191; PMID:2501755
A;Accession: S04722
A;Molecule type: DNA
A;Residues: 1-883 <JAN>
A;Cross-references: UNIPROT:P11536; UNIPARC:UPI0000129AFF; EMBL:X15087; NID:g7513; PIDN:
A;Note: Gln-867 was also found
R;Burtis, K.C.; Thummel, C.S.; Jones, C.W.; Karim, F.D.; Hogness, D.S.
Cell 61, 85-99, 1990
A;Title: The Drosophila 74EF early puff contains E74, a complex ecdysone-inducible gene
A;Reference number: A90912; MUID:90199900; PMID:2107982
A;Accession: B34692
A;Molecule type: DNA
A;Residues: 1-866,'Q',868-883 <BUR>
A;Cross-references: UNIPARC:UPI00001247C8; GB:M37083; NID:g157309; PID:g157310
C;Genetics:
A;Gene: FlyBase:Eip74EF
A;Cross-references: FlyBase:FBgn0000567
A;Map position: 3 74EF
C;Keywords: alternative splicing; DNA binding; transcription regulation
F;789-869/Domain: ets DNA-binding domain homology <ETS>

```
  Query Match             31.6%;  Score 56;  DB 2;  Length 883;
  Best Local Similarity   60.0%;  Pred. No. 21;
  Matches   15;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          3 SSLFSPSSSSYSSLFTAKPTRLLSP 27
              |||   |||| ||| :| || : ||
Db         39 SSLSLSSSSSSSSLSSATPTPVASP 63
```

RESULT 9
T48103
mRNA binding protein CSP41 homolog T20O10.240 [similarity] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T48103
R;Obermaier, B.; Ottenwaelder, B.; Duchemin, D.; Zeitler, K.; Mewes, H.W.; Rudd, S.; Lemc
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24484
A;Accession: T48103
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-406 <OBE>
A;Cross-references: UNIPROT:Q9LYA9; UNIPARC:UPI000009DE02; EMBL:AL163816
A;Experimental source: cultivar Columbia; BAC clone T20O10
C;Genetics:
A;Map position: 3
A;Introns: 132/3; 173/2; 207/3; 254/1; 346/3
A;Note: T20O10.240

```
  Query Match             31.1%;  Score 55;  DB 2;  Length 406;
  Best Local Similarity   58.1%;  Pred. No. 12;
  Matches   18;  Conservative    4;  Mismatches    7;  Indels    2;  Gaps    2;

Qy          3 SSLFSPSSSSYSSL-FTAKPTRLLSPKPKFT 32
              ||| | |||| | | |: : :| |||: |||
Db         41 SSLSSSSSSSSSLLTFSLRTSRRLSPQ-KFT 70
```

RESULT 10
T38714
hypothetical protein SPAC3F10.13 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T38714
R;Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, February 1996
A;Reference number: Z21807
A;Accession: T38714
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-612 <MUR>
A;Cross-references: UNIPROT:Q10187; UNIPARC:UPI000013A18B; EMBL:Z69369; PIDN:CAA93311.1;
A;Experimental source: strain 972h-; cosmid c3F10
C;Genetics:
A;Gene: SPDB:SPAC3F10.13
A;Map position: 1
C;Superfamily: Schizosaccharomyces pombe hypothetical protein SPAC3F10.13

```
  Query Match             30.8%;  Score 54.5;  DB 2;  Length 612;
  Best Local Similarity   32.6%;  Pred. No. 22;
  Matches   14;  Conservative    7;  Mismatches   13;  Indels    9;  Gaps    1;

Qy          3 SSLFSPSSSSYSSLF---------TAKPTRLLSPKPKFTFSIR 36
              || ||| :: : | |           : ||| : : ||  |  :|
Db        234 SSYFSPPNAEFDSTFCSTLSSLTNSEKPTYIFNFKPNSTPDLR 276
```

RESULT 11
T30200
protein-tyrosine kinase (EC 2.7.1.112) alk - mouse
N;Alternate names: anaplastic lymphoma kinase
C;Species: Mus musculus (house mouse)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C;Accession: T30200
R;Iwahara, T.; Fujimoto, J.; Wen, D.; Cupples, R.; Bucay, N.; Arakawa, T.; Mori, S.; Rat
Oncogene 14, 439-449, 1997
A;Title: Molecular characterization of ALK, a receptor tyrosine kinase expressed specific
A;Reference number: Z20774; MUID:97178863; PMID:9053841
A;Accession: T30200
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA

```
A;Residues: 1-1621 <IWA>
A;Cross-references: UNIPROT:P97793; UNIPARC:UPI0000029516; EMBL:D83002; NID:g1864006; PI
A;Experimental source: brain and testis
C;Genetics:
A;Gene: alk
C;Function:
A;Description: may play an important role in development of the brain
C;Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase

  Query Match             30.8%;  Score 54.5;  DB 2;  Length 1621;
  Best Local Similarity   44.4%;  Pred. No. 65;
  Matches   12;  Conservative    5;  Mismatches    7;  Indels    3;  Gaps    1;

Qy          3 SSLFSPSSSSYSSLFTAKPTRLLSPKP 29
              :||::|   :| | ||||| :   | |
Db       1509 TSLWNP---TYGSWFTAKPAKKTHPPP 1532


RESULT 12
A75491
hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: A75491
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: A75491
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-132 <WHI>
A;Cross-references: UNIPROT:Q9RWL1; UNIPARC:UPI00000C17C8; GB:AE001923; GB:AE000513; NID
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0655
A;Map position: 1
C;Superfamily: Deinococcus radiodurans hypothetical protein DR0655

  Query Match             30.5%;  Score 54;  DB 2;  Length 132;
  Best Local Similarity   48.4%;  Pred. No. 4.6;
  Matches   15;  Conservative    4;  Mismatches   10;  Indels    2;  Gaps    1;

Qy          3 SSLFSPSSSS--YSSLFTAKPTRLLSPKPKF 31
              |:|  |:|||  |||| :| | |:  |   |
Db          4 SALAQPASSSVGYSSLVSAVPQRVTLPLQLF 34


RESULT 13
T47591
aintegumaenta-like protein - Arabidopsis thaliana
N;Alternate names: protein T12E18.10
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T47591
R;Bloecker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat M.Mewes,
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24469
A;Accession: T47591
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-205 <BLO>
A;Cross-references: UNIPROT:Q9M2V4; UNIPARC:UPI00000AAF3C; EMBL:AL132971
A;Experimental source: cultivar Columbia; BAC clone T12E18
C;Genetics:
A;Map position: 3
A;Introns: 71/2; 99/1; 128/3; 170/2
A;Note: T12E18.10

  Query Match             30.5%;  Score 54;  DB 2;  Length 205;
  Best Local Similarity   50.0%;  Pred. No. 7.6;
  Matches   12;  Conservative    4;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          7 SPSSSSYSSLFTAKPTRLLSPKPK 30
              |||||  ||  |: | :  :|:||
Db         13 SPSSSVSSSTTTSSPIQSEAPRPK 36


RESULT 14
T45712
NAD-dependent malate dehydrogenase, chloroplast - Arabidopsis thaliana
N;Alternate names: protein F1P2.70
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C;Accession: T45712
R;Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sau
submitted to the Protein Sequence Database, November 1999
A;Reference number: Z23010
A;Accession: T45712
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-403 <CHO>
A;Cross-references: UNIPROT:Q9SN86; UNIPARC:UPI00000AC228; EMBL:AL132955
A;Experimental source: cultivar Columbia; BAC clone F1P2
C;Genetics:
A;Map position: 3
A;Genome: nuclear
A;Note: F1P2.70
C;Superfamily: L-lactate dehydrogenase
C;Keywords: chloroplast

  Query Match             30.5%;  Score 54;  DB 2;  Length 403;
  Best Local Similarity   47.1%;  Pred. No. 16;
  Matches   16;  Conservative    3;  Mismatches   11;  Indels    4;  Gaps    1;

Qy          3 SSLFSPSSSSYSSLFTAKPTRLLSPK----PKFT 32
              :|||| |||||   :   :|| | |    | ||
Db          7 ASLFSTVSSSYSKASSIPHSRLQSVKFNSVPSFT 40


RESULT 15
T51862
malate dehydrogenase (EC 1.1.1.37), chloroplast [validated] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C;Accession: T51862
R;Berkemeyer, M.; Scheibe, R.; Ocheretina, O.
J. Biochem. 273, 27927-27933, 1998
A;Title: A novel, non-redox-regulated NAD-dependent malate dehydrogenase from chloroplast
A;Reference number: Z25488
A;Accession: T51862
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-403 <BER>
A;Cross-references: UNIPROT:O81844; UNIPARC:UPI00000A9B2A; EMBL:Y13987; PIDN:CAA74320.1
C;Genetics:
A;Gene: mdh
A;Genome: nuclear
C;Function:
A;Description: EC 1.1.1.37 [validated, MUID:98447636]
A;Note: NAD-dependent
C;Superfamily: L-lactate dehydrogenase
C;Keywords: chloroplast; NAD; oxidoreductase

  Query Match             30.5%;  Score 54;  DB 2;  Length 403;
  Best Local Similarity   47.1%;  Pred. No. 16;
  Matches   16;  Conservative    3;  Mismatches   11;  Indels    4;  Gaps    1;

Qy          3 SSLFSPSSSSYSSLFTAKPTRLLSPK----PKFT 32
              :|||| |||||   :   :|| | |    | ||
Db          7 ASLFSTVSSSYSKASSIPHSRLQSVKFNSVPSFT 40


Search completed: March  7, 2006, 21:59:04
```

Job time : 30.8 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2006, 21:43:16 ; Search time 178.2 Seconds
(without alignments)
142.531 Million cell updates/sec

Title: US-10-751-235-11
Perfect score: 177
Sequence: 1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	177	100.0	539	2	Q6TBX7_ARATH	Q6tbx7 arabidopsis
2	177	100.0	552	2	Q8RWV4_ARATH	Q8rwv4 arabidopsis
3	177	100.0	566	2	Q9SCP8_ARATH	Q9scp8 arabidopsis
4	65.5	37.0	430	2	Q9URW6_SCHPO	Q9urw6 schizosacch
5	62.5	35.3	1332	2	Q8MQQ8_DROME	Q8mqq8 drosophila
6	62.5	35.3	2016	2	Q9VLS7_DROME	Q9vls7 drosophila
7	61.5	34.7	1383	2	Q5B1U5_EMENI	Q5b1u5 aspergillus
8	60	33.9	389	1	PUR5_ARATH	Q05728 arabidopsis
9	59.5	33.6	319	1	RCCR_ARATH	Q81du4 arabidopsis
10	59	33.3	624	2	Q7S9I0_NEUCR	Q7s9i0 neurospora
11	59	33.3	1545	2	Q9VYN1_DROME	Q9vyn1 drosophila
12	59	33.3	1733	2	Q4H4D7_MOUSE	Q4h4d7 mus musculu
13	58.5	33.1	355	2	Q17793_CAEEL	Q17793 caenorhabdi
14	58	32.8	605	2	Q75WN1_CHAOB	Q75wn1 chamaecypar
15	58	32.8	771	2	Q86H13_DICDI	Q86h13 dictyosteli
16	58	32.8	2408	2	Q57UP3_9TRYP	Q57up3 trypanosoma
17	57.5	32.5	512	2	Q55P26_CRYNE	Q55p26 cryptococcu
18	57.5	32.5	513	2	Q5KED0_CRYNE	Q5ked0 cryptococcu
19	57	32.2	126	2	Q693I6_CRYHO	Q693i6 cryptospori
20	57	32.2	144	2	Q4XCA2_PLACH	Q4xca2 plasmodium
21	57	32.2	302	2	Q9C589_ARATH	Q9c589 arabidopsis
22	57	32.2	1449	2	Q86A82_DICDI	Q86a82 dictyosteli
23	57	32.2	1464	2	Q555X3_DICDI	Q555x3 dictyosteli
24	56.5	31.9	497	2	Q6MAZ9_PARUW	Q6maz9 parachlamyd
25	56.5	31.9	577	2	Q620Q0_CAEBR	Q620q0 caenorhabdi
26	56	31.6	353	2	Q8X0N1_NEUCR	Q8x0n1 neurospora
27	56	31.6	357	2	Q9D347_MOUSE	Q9d347 mus musculu
28	56	31.6	434	2	Q8LED3_ARATH	Q8led3 arabidopsis
29	56	31.6	459	2	Q5F245_MOUSE	Q5f245 mus musculu
30	56	31.6	474	2	Q9M2K7_ARATH	Q9m2k7 arabidopsis
31	56	31.6	476	2	Q8W4L8_ARATH	Q8w4l8 arabidopsis
32	56	31.6	497	2	Q6PDE7_MOUSE	Q6pde7 mus musculu
33	56	31.6	883	1	E74EB_DROME	P11536 drosophila
34	56	31.6	1061	2	Q555D2_DICDI	Q555d2 dictyosteli
35	56	31.6	1806	2	Q869R4_DICDI	Q869r4 dictyosteli
36	55	31.1	143	2	Q8LA28_ARATH	Q8la28 arabidopsis
37	55	31.1	406	2	Q9LYA9_ARATH	Q9lya9 arabidopsis
38	55	31.1	818	2	Q6J735_SORBI	Q6j735 sorghum bic
39	55	31.1	836	2	Q4RM98_TETNG	Q4rm98 tetraodon n
40	54.5	30.8	138	2	Q6ZFU9_ORYSA	Q6zfu9 oryza sativ
41	54.5	30.8	416	2	Q6Z8R9_ORYSA	Q6z8r9 oryza sativ
42	54.5	30.8	612	1	UCP6_SCHPO	Q10187 schizosacch
43	54.5	30.8	696	2	Q86ZL2_PODAN	Q86zl2 podospora a
44	54.5	30.8	1621	1	ALK_MOUSE	P97793 mus musculu
45	54.5	30.8	2027	1	DOCK3_MOUSE	Q8ciq7 mus musculu

ALIGNMENTS

RESULT 1
Q6TBX7_ARATH
ID Q6TBX7_ARATH PRELIMINARY; PRT; 539 AA.
AC Q6TBX7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Chloroplast carotenoid epsilon-ring hydroxylase.
GN Name=LUT1;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14709673; DOI=10.1073/pnas.2237237100;
RA Tian L., Musetti V., Kim J., Magallanes-Lundback M., DellaPenna D.;
RT "The Arabidopsis LUT1 locus encodes a member of the cytochrome p450
RT family that is required for carotenoid epsilon-ring hydroxylation
RT activity.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:402-407(2004).
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; AY424805; AAR83120.1; -; mRNA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002401; EP450I.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR00463; EP450I.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase;
KW Transmembrane.
SQ SEQUENCE 539 AA; 60555 MW; 4C25C728B676AEEB CRC64;

Query Match 100.0%; Score 177; DB 2; Length 539;
Best Local Similarity 100.0%; Pred. No. 9.8e-15;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36
||||||||||||||||||||||||||||||||||||
Db 1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36

RESULT 2
Q8RWV4_ARATH
ID Q8RWV4_ARATH PRELIMINARY; PRT; 552 AA.
AC Q8RWV4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative cytochrome P450 (Fragment).

```
GN   Name=At3g53130;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC   rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA   Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA   Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA   Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA   Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA   Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA   Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.;
RA   Davis R.W., Ecker J.R., Theologis A.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR   EMBL; AY091083; AAM13903.1; -; mRNA.
DR   HSSP; P14779; 1JPZ.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   InterPro; IPR002401; EP450I.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00463; EP450I.
DR   PRINTS; PR00385; P450.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase;
KW   Transmembrane.
FT   NON_TER       1       1
SQ   SEQUENCE   552 AA;  62073 MW;  C23CF8498B5B8440 CRC64;

  Query Match             100.0%;  Score 177;  DB 2;  Length 552;
  Best Local Similarity   100.0%;  Pred. No. 1e-14;
  Matches   36;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36
              |||||||||||||||||||||||||||||||||||
Db         14 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 49


RESULT 3
Q9SCP8_ARATH
ID   Q9SCP8_ARATH PRELIMINARY;      PRT;   566 AA.
AC   Q9SCP8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Cytochrom P450-like protein.
GN   Name=T4D2.60;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC   rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Nyakatura G., Fartmann B., Dauner D., Sterr W., Holland R.,
RA   Weichselgartner M., Mewes H.W., Lemcke K., Mayer K.F.X., Quetier F.,
RA   Salanoubat M.;
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   EU Arabidopsis sequencing project;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR   EMBL; AL132958; CAB64216.1; -; Genomic_DNA.
DR   PIR; T46159; T46159.
DR   HSSP; P14779; 1JPZ.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   InterPro; IPR002401; EP450I.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00463; EP450I.
DR   PRINTS; PR00385; P450.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase;
KW   Transmembrane.
SQ   SEQUENCE   566 AA;  63596 MW;  B884E8996B1A4C7D CRC64;

  Query Match             100.0%;  Score 177;  DB 2;  Length 566;
  Best Local Similarity   100.0%;  Pred. No. 1e-14;
  Matches   36;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36
              |||||||||||||||||||||||||||||||||||
Db          1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36


RESULT 4
Q9URW6_SCHPO
ID   Q9URW6_SCHPO PRELIMINARY;      PRT;   430 AA.
AC   Q9URW6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   SPAPJ696.02 protein.
GN   ORFNames=SPAPJ696.02;
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M.H., Lyne R., Stewart A.,
RA   Sgouros J.G., Peat N., Hayles J., Baker S.G., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D.E., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K.D., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K.L., Murphy L.D., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K.M., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M.N., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J.R., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Duesterhoeft A., Fritzc C., Holzer E., Moestl D.,
RA   Hilbert H., Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R.,
RA   Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
CC   -!- SIMILARITY: Contains 1 SH3 domain.
DR   EMBL; AL133359; CAB62422.1; -; Genomic_DNA.
DR   PIR; T50296; T50296.
DR   HSSP; P08631; 1BU1.
DR   GeneDB_Spombe; SPAPJ696.02; -.
DR   InterPro; IPR007461; DUF500.
DR   InterPro; IPR001452; SH3.
DR   Pfam; PF04366; DUP500; 1.
DR   Pfam; PF00018; SH3_1; 1.
```

```
DR   PRINTS; PR00452; SH3DOMAIN.
DR   ProDom; PD000066; SH3; 1.
DR   SMART; SM00326; SH3; 1.
DR   PROSITE; PS50002; SH3; 1.
KW   Complete proteome; SH3 domain.
SQ   SEQUENCE   430 AA;  46373 MW;  AC300C08C29D946F CRC64;

  Query Match             37.0%;  Score 65.5;  DB 2;  Length 430;
  Best Local Similarity   55.2%;  Pred. No. 4.4;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    5;  Gaps    1;

Qy          3 SSLFSPSSSSYSSLFTAKPTRLLSPKPKF 31
              || ||  || ||     ||:|  :|||||
Db        343 SSQFSSRSSSEYS-----KPSRPTAPKPKF 366


RESULT 5
Q8MQQ8_DROME
ID   Q8MQQ8_DROME PRELIMINARY;      PRT;  1332 AA.
AC   Q8MQQ8;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   LD21041p (Fragment).
GN   ORFNames=CG8552;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA   Celniker S.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY128444; AAM75037.1; -; mRNA.
DR   Ensembl; CG8552; Drosophila melanogaster.
DR   FlyBase; FBgn0031990; CG8552.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   InterPro; IPR004177; DDHD.
DR   InterPro; IPR004170; WWE.
DR   Pfam; PF02862; DDHD; 1.
DR   Pfam; PF02825; WWE; 1.
DR   PROSITE; PS50918; WWE; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   1332 AA;  144944 MW;  B3F3F5931D907008 CRC64;

  Query Match             35.3%;  Score 62.5;  DB 2;  Length 1332;
  Best Local Similarity   47.4%;  Pred. No. 41;
  Matches   18;  Conservative    4;  Mismatches   11;  Indels    5;  Gaps    2;

Qy          3 SSLFSPSS-SSYSSLFTAKP----TRLLSPKPKFTFSI 35
              |||||||   | ||||   |    | :|:| |  | ::
Db        302 SSLFSPSQVDSTSSLFAPPPTSANTTILTPAPDTTSTL 339


RESULT 6
Q9VLS7_DROME
ID   Q9VLS7_DROME PRELIMINARY;      PRT;  2016 AA.
AC   Q9VLS7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   CG8552-PA.
GN   Name=CG8552; ORFNames=CG8552;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426065; PubMed=12537568;
RA   Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA   Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA   George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA   Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA   Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA   Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT   "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT   melanogaster euchromatic genome sequence.";
RL   Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426070; PubMed=12537573;
RA   Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA   Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA   Ashburner M., Celniker S.E.;
RT   "The transposable elements of the Drosophila melanogaster euchromatin:
RT   a genomics perspective.";
RL   Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
```

```
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN   [5]
RP   NUCLEOTIDE SEQUENCE.
RG   Berkeley Drosophila Genome Project;
RA   Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA   Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA   Yu C., Rubin G.;
RT   "Drosophila melanogaster release 4 sequence.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   NUCLEOTIDE SEQUENCE.
RG   FlyBase;
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
CC   -!- INTERACTION:
CC       Q9VCB1:CG13609; NbExp=1; IntAct=EBI-190906, EBI-105274;
DR   EMBL; AE003620; AAF52607.1; -; Genomic_DNA.
DR   IntAct; Q9VLS7; -.
DR   Ensembl; CG8552; Drosophila melanogaster.
DR   FlyBase; FBgn0031990; CG8552.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   InterPro; IPR004177; DDHD.
DR   InterPro; IPR004170; WWE.
DR   Pfam; PF02862; DDHD; 1.
DR   Pfam; PF02825; WWE; 1.
DR   PROSITE; PS50918; WWE; 1.
SQ   SEQUENCE   2016 AA;  214155 MW;  05C36B985F268641 CRC64;

  Query Match             35.3%;  Score 62.5;  DB 2;  Length 2016;
  Best Local Similarity   47.4%;  Pred. No. 67;
  Matches   18;  Conservative    4;  Mismatches   11;  Indels    5;  Gaps    2;

Qy          3 SSLFSPSS-SSYSSLFTAKP----TRLLSPKPKFTFSI 35
              |||||||   | ||||   |    | :|:| |  | ::
Db        986 SSLFSPSQVDSTSSLFAPPPTSANTTILTPAPDTTSTL 1023


RESULT 7
Q5B1U5_EMENI
ID   Q5B1U5_EMENI PRELIMINARY;      PRT;  1383 AA.
AC   Q5B1U5;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=AN5485.2;
OS   Aspergillus nidulans FGSC A4.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC   Eurotiales; Trichocomaceae; Emericella.
OX   NCBI_TaxID=227321;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=FGSC A4;
RA   Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA   Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA   Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA   Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA   Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA   Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA   Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA   Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA   Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA   Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA   Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA   Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA   Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA   Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA   Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA   Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA   Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA   Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA   Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA   Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA   Lander E.;
RT   "Genome Sequence of Aspergillus nidulans.";
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AACD01000094; EAA62645.1; -; Genomic_DNA.
DR   InterPro; IPR011993; PH_type.
DR   InterPro; IPR000156; Ran_BP1.
DR   Pfam; PF00638; Ran_BP1; 1.
DR   SMART; SM00160; RanBD; 1.
DR   PROSITE; PS50196; RANBD1; 1.
KW   Hypothetical protein; Nuclear protein; Protein transport; Transport.
SQ   SEQUENCE   1383 AA;  141244 MW;  D7F854C0423511E7 CRC64;

  Query Match             34.7%;  Score 61.5;  DB 2;  Length 1383;
  Best Local Similarity   37.5%;  Pred. No. 58;
  Matches   15;  Conservative    4;  Mismatches    8;  Indels   13;  Gaps    1;

Qy          4 SLFSPSSSS----------YSSLFTAKPTRLLSPKPK 30
              |||:|  :|           :|||| ||   ||||:
Db        431 SLFAPEPASTEQEKPKPAEGNPFSSLFVPKPANAESPKPE 470


RESULT 8
PUR5_ARATH
ID   PUR5_ARATH     STANDARD;      PRT;   389 AA.
AC   Q05728; Q9M2W5; Q9SV49;
DT   01-NOV-1995 (Rel. 32, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Phosphoribosylformylglycinamidine cyclo-ligase, chloroplast precursor
DE   (EC 6.3.3.1) (AIRS) (Phosphoribosyl-aminoimidazole synthetase) (AIR
DE   synthase).
GN   Name=PUR5; OrderedLocusNames=At3g55010; ORFNames=T15C9.10, F28P10.10;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC   rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=cv. Columbia; TISSUE=Leaf, and Stem;
RX   MEDLINE=94151435; PubMed=8108507; DOI=10.1104/pp.102.2.387;
RA   Senecoff J.F., Meagher R.B.;
RT   "Isolating the Arabidopsis thaliana genes for de novo purine synthesis
RT   by suppression of Escherichia coli mutants. I. 5'-Phosphoribosyl-5-
RT   aminoimidazole synthetase.";
RL   Plant Physiol. 102:387-399(1993).
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=cv. Columbia;
RX   MEDLINE=21016720; PubMed=11130713; DOI=10.1038/35048706;
RA   Salanoubat M., Lemcke K., Rieger M., Ansorge W., Unseld M.,
RA   Fartmann B., Valle G., Bloecker H., Perez-Alonso M., Obermaier B.,
RA   Delseny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
RA   De Simone V., Choisne N., Artiguenave F., Robert C., Brottier P.,
RA   Wincker P., Cattolico L., Weissenbach J., Saurin W., Quetier F.,
RA   Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA   Wurmbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,
RA   Wiedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA   Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simionati B.,
RA   Conrad A., Hornischer K., Kauer G., Loehnert T.-H., Nordsiek G.,
RA   Reichelt J., Scharfe M., Schoen O., Bargues M., Terol J., Climent J.,
RA   Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA   Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,
RA   de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA   Monfort A., Argiriou A., Flores M., Liguori R., Vitale D.,
RA   Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA   Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA   Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,
RA   Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA   Pai G., Militscher J., Sellers P., Gill J.E., Feldblyum T.V.,
```

```
RA   Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA   Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA   Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA   Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA   Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
RA   Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT   "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 408:820-822(2000).
RN   [3]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   STRAIN=cv. Columbia;
RX   MEDLINE=22954850; PubMed=14593172; DOI=10.1126/science.1088305;
RA   Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
RA   Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA   Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA   Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA   Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA   Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA   Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
RA   Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
RA   Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA   Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA   Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
RA   Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT   "Empirical analysis of transcriptional activity in the Arabidopsis
RT   genome.";
RL   Science 302:842-846(2003).
CC   -!- CATALYTIC ACTIVITY: ATP + 2-(formamido)-N(1)-(5-phospho-D-
CC       ribosyl)acetamidine = ADP + phosphate + 5-amino-1-(5-phospho-D-
CC       ribosyl)imidazole.
CC   -!- PATHWAY: Nucleotide biosynthesis; IMP biosynthesis; 5-amino-1-(5-
CC       phospho-D-ribosyl)imidazole-4-carboxamide from N(2)-formyl-N(1)-
CC       (5-phospho-D-ribosyl)glycinamide: step 2.
CC   -!- SUBCELLULAR LOCATION: Chloroplast.
CC   -!- SIMILARITY: Belongs to the AIR synthase family.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC       frameshift in position 344.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; L12457; AAC37341.1; ALT_FRAME; Unassigned_DNA.
DR   EMBL; AL132970; CAB82696.1; -; Genomic_DNA.
DR   EMBL; AL049655; CAB41083.1; -; Genomic_DNA.
DR   EMBL; AY060585; AAL31210.1; -; mRNA.
DR   EMBL; AY142054; AAM98318.1; -; mRNA.
DR   PIR; T47640; T47640.
DR   HSSP; P08178; 1CLI.
DR   InterPro; IPR000728; AIR_synth.
DR   InterPro; IPR010918; AIR_synth_C.
DR   InterPro; IPR004733; PurM_cligase.
DR   Pfam; PF00586; AIRS; 1.
DR   Pfam; PF02769; AIRS_C; 1.
DR   TIGRFAMs; TIGR00878; purM; 1.
KW   Chloroplast; Ligase; Purine biosynthesis; Transit peptide.
FT   TRANSIT      1     58       Chloroplast (Potential).
FT   CHAIN       59    389       Phosphoribosylformylglycinamidine cyclo-
FT                               ligase.
SQ   SEQUENCE   389 AA;  41504 MW;  F341E9D21E9286F0 CRC64;

  Query Match             33.9%;  Score 60;  DB 1;  Length 389;
  Best Local Similarity   42.1%;  Pred. No. 21;
  Matches   16;  Conservative    6;  Mismatches   10;  Indels    6;  Gaps    2;

Qy          1 MESSLFSPSSSSYSSLFTAKPTR---LLSPKPKFTPSI 35
              ||: :   ||| ||||:   :|   : ||||   ||:
Db          1 MEARILQSSSSCYSSLYAVNRSRFSSVSSPKP---FSV 35
```

```
RESULT 9
RCCR_ARATH
ID   RCCR_ARATH     STANDARD;      PRT;   319 AA.
AC   Q8LDU4; O23185;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Red chlorophyll catabolite reductase, chloroplast precursor
DE   (EC 1.-.-.-) (RCC reductase) (AtRCCR) (Accelerated cell death protein
DE   2).
GN   Name=RCCR; Synonyms=ACD2; OrderedLocusNames=At4g37000;
GN   ORFNames=C7A10_360;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC   rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   NUCLEOTIDE SEQUENCE, FUNCTION, AND SUBCELLULAR LOCATION.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20208062; PubMed=10743659;
RA   Wuethrich K.L., Bovet L., Hunziker P.E., Donnison I.S.,
RA   Hoertensteiner S.;
RT   "Molecular cloning, functional expression and characterization of RCC
RT   reductase, involved in chlorophyll catabolism.";
RL   Plant J. 21:189-198(2000).
RN   [2]
RP   NUCLEOTIDE SEQUENCE, SUBCELLULAR LOCATION, AND MUTANTS ACD2-7;
RP   ACD2-12E13 AND ACD2-6/ACD2-8.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=21143384; PubMed=11149948; DOI=10.1073/pnas.021465298;
RA   Mach J.M., Castillo A.R., Hoogstraten R., Greenberg J.T.;
RT   "The Arabidopsis accelerated cell death gene ACD2 encodes red
RT   chlorophyll catabolite reductase and suppresses the spread of disease
RT   symptoms.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:771-776(2001).
RN   [3]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083488; PubMed=10617198; DOI=10.1038/47134;
RA   Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA   Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA   Harris B., Ansorge W., Brandt P., Grivell L.A., Rieger M.,
RA   Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA   Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
RA   Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA   Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA   Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA   Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA   Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA   Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA   Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA   Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA   Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA   De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
RA   Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,
RA   Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
RA   Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA   Borkova D., Bloecker H., Scharfe M., Grimm M., Loehnert T.-H.,
RA   Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA   Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
RA   Massenet O., Quigley F., Clabauld G., Muendlein A., Felber R.,
RA   Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA   Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA   Gibbons T., Weber N., Vandenbol M., Bargues M., Terol J., Torres A.,
RA   Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA   Heijnen L., Schwarz S., Scholler P., Heber S., Francs P., Bielke C.,
RA   Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA   Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA   Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA   Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA   Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
```

```
RA   Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA   Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA   Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.W.,
RA   Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA   Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshu C.,
RA   Antonoiu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA   Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA   Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA   Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA   Chen E., Marra M.A., Martienssen R., McCombie W.R.;
RT   "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:769-777(1999).
RN   [4]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   STRAIN=cv. Columbia;
RX   MEDLINE=22954850; PubMed=14593172; DOI=10.1126/science.1088305;
RA   Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
RA   Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA   Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA   Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA   Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA   Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA   Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
RA   Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
RA   Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA   Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA   Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
RA   Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT   "Empirical analysis of transcriptional activity in the Arabidopsis
RT   genome.";
RL   Science 302:842-846(2003).
RN   [5]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RA   Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA   Peldmann K.A.;
RT   "Full-length cDNA from Arabidopsis thaliana.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Catalyzes the key reaction of chlorophyll catabolism,
CC       porphyrin macrocycle cleavage of pheophorbide a (pheide a) to a
CC       primary fluorescent catabolite (pFCC). Works in a two-step
CC       reaction with pheophorbide a oxygenase (PaO) by reducing the
CC       C20/C1 double bond of the intermediate, RCC.
CC   -!- COFACTOR: Reduced ferredoxin.
CC   -!- PATHWAY: Chlorophyll degradation; fourth step.
CC   -!- SUBCELLULAR LOCATION: Chloroplast stroma, and a low amount in
CC       mitochondria of 7-day-old seedlings.
CC   -!- TISSUE SPECIFICITY: Expressed in all tissues tested, including
CC       roots.
CC   -!- DEVELOPMENTAL STAGE: Present at all times of development. No
CC       change of levels during senescence or pathogen attack.
CC   -!- MISCELLANEOUS: The absence of light completely suppresses cell
CC       death in acd2 mutants.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AF326347; AAG53980.1; -; mRNA.
DR   EMBL; Z99707; CAB16763.1; -; Genomic_DNA.
DR   EMBL; AL161590; CAB80366.1; -; Genomic_DNA.
DR   EMBL; AY045578; AAK73936.1; -; mRNA.
DR   EMBL; AY093785; AAM10401.1; -; mRNA.
DR   EMBL; AY085797; AAM63013.1; -; mRNA.
DR   InterPro; IPR009439; RCC_reductase.
DR   Pfam; PF06405; RCC_reductase; 1.
KW   Chlorophyll catabolism; Chloroplast; Coiled coil; Oxidoreductase;
KW   Transit peptide.
FT   TRANSIT       1     39       Chloroplast (Potential).
FT   CHAIN        40    319       Red chlorophyll catabolite reductase.
FT   COILED      255    286       Potential.
FT   MUTAGEN     140    140       G->V: In acd2-12E13; spontaneous
FT                                spreading cell death lesions.
FT   MUTAGEN     181    192       Missing: In acd2-7; spontaneous spreading
FT                                cell death lesions.
FT   MUTAGEN     279    279       R->K: In acd2-6; spontaneous spreading
FT                                cell death lesions.
FT   CONFLICT    290    290       D -> E (in Ref. 5).
SQ   SEQUENCE   319 AA;  36449 MW;  A46DC65FB7452517 CRC64;

  Query Match             33.6%;  Score 59.5;  DB 1;  Length 319;
  Best Local Similarity   43.8%;  Pred. No. 19;
  Matches   14;  Conservative    7;  Mismatches    8;  Indels    3;  Gaps    1;

Qy          3 SSLFSPSSSSYSSLFTAKPTRL---LSPKPKF 31
              ::|:| || || |  |:||:|    | |: :|
Db          7 NTLYSSSSPSYLSPLTSKPSRFSKNLRPRAQF 38


RESULT 10
Q7S9I0_NEUCR
ID   Q7S9I0_NEUCR PRELIMINARY;      PRT;   624 AA.
AC   Q7S9I0;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Predicted protein.
GN   Name=NCU06547.1;
OS   Neurospora crassa.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX   NCBI_TaxID=5141;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=OR74A;
RA   Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA   Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA   Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA   Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA   Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA   Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA   Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA   Kamal M., Kamvysselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA   Krystofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,
RA   Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmani S.A.,
RA   DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA   Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA   Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA   Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT   "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL   Nature 0:0-0(2003).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AABX01000213; EAA33018.1; -; Genomic_DNA.
SQ   SEQUENCE   624 AA;  65243 MW;  54F90D3DC736005C CRC64;

  Query Match             33.3%;  Score 59;  DB 2;  Length 624;
  Best Local Similarity   51.9%;  Pred. No. 49;
  Matches   14;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          3 SSLFSPSSSSYSSLFTAKPTRLLSPKP 29
              ||  | |||| || |:: |: | :| |
Db        287 SSSSSSSSSSSSSAFSSLPSSLTTPSP 313


RESULT 11
Q9VYN1_DROME
ID   Q9VYN1_DROME PRELIMINARY;      PRT;  1545 AA.
AC   Q9VYN1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   CG11245-PA.
```

```
GN   Name=CG11245; ORFNames=CG11245;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426065; PubMed=12537568;
RA   Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA   Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA   George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA   Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA   Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA   Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT   "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT   melanogaster euchromatic genome sequence.";
RL   Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426070; PubMed=12537573;
RA   Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA   Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA   Ashburner M., Celniker S.E.;
RT   "The transposable elements of the Drosophila melanogaster euchromatin:
RT   a genomics perspective.";
RL   Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN   [5]
RP   NUCLEOTIDE SEQUENCE.
RG   Berkeley Drosophila Genome Project;
RA   Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA   Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA   Yu C., Rubin G.;
RT   "Drosophila melanogaster release 4 sequence.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   NUCLEOTIDE SEQUENCE.
RG   FlyBase;
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003489; AAF48161.1; -; Genomic_DNA.
DR   Ensembl; CG11245; Drosophila melanogaster.
DR   FlyBase; FBgn0030388; CG11245.
SQ   SEQUENCE   1545 AA;  166082 MW;  2ABCEABED136C8B8 CRC64;

  Query Match             33.3%;  Score 59;  DB 2;  Length 1545;
  Best Local Similarity   44.1%;  Pred. No. 1.4e+02;
  Matches   15;  Conservative    3;  Mismatches   12;  Indels    4;  Gaps    1;

Qy          2 ESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSI 35
              || |  | :|   |  |  |     ||||||| ::
Db        475 ESLLIQPLTSGSGSSSTPTP----SPKPKFTVNV 504


RESULT 12
Q4H4D7_MOUSE
ID   Q4H4D7_MOUSE PRELIMINARY;      PRT;  1733 AA.
AC   Q4H4D7;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   ATP-binding cassette (ABC) transporter ABCA14.
GN   Name=ABCA14;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6N; TISSUE=Testis;
RA   Ban N., Sasaki M., Sakai H., Ueda K., Inagaki N.;
RT   "Cloning of ABCA17, a novel rodent sperm-specific ABC (ATP-binding
RT   cassette) transporter that regulates intracellular lipid metabolism.";
RL   Biochem. J. 389:577-585(2005).
DR   EMBL; AB112584; BAD97416.1; -; mRNA.
KW   ATP-binding.
SQ   SEQUENCE   1733 AA;  195943 MW;  2D4DB5B8D19F6D6E CRC64;

  Query Match             33.3%;  Score 59;  DB 2;  Length 1733;
  Best Local Similarity   44.8%;  Pred. No. 1.6e+02;
  Matches   13;  Conservative    6;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          2 ESSLFSPSSSSYSSLFTAKPTRLLSPKPK 30
              :::| |||  | |   :: |:|| || ||
Db       1692 QAALASPSPPSNSRPISSPPSRLSSPTPK 1720


RESULT 13
Q17793_CAEEL
ID   Q17793_CAEEL PRELIMINARY;      PRT;   355 AA.
AC   Q17793; Q22365;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
```

```
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein T09F3.1.
GN   ORFNames=T09F3.1;
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916; DOI=10.1126/science.282.5396.2012;
RG   The C. elegans sequencing consortium;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
CC   -!- INTERACTION:
CC       Q9TZ38:gei-4; NbExp=1; IntAct=EBI-329146, EBI-329192;
CC       Q9U9Y8:lit-1; NbExp=1; IntAct=EBI-329146, EBI-318513;
DR   EMBL; Z49908; CAA90104.2; -; Genomic_DNA.
DR   EMBL; Z49070; CAA88872.2; -; Genomic_DNA.
DR   EMBL; Z49070; CAA90104.2; JOINED; Genomic_DNA.
DR   EMBL; Z49908; CAA88872.2; JOINED; Genomic_DNA.
DR   PIR; T19059; T19059.
DR   IntAct; Q17793; -.
DR   Ensembl; T09F3.1; Caenorhabditis elegans.
DR   WormBase; WBGene00011661; T09F3.1.
DR   WormPep; T09F3.1; CE35906.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0007275; P:development; IEA.
DR   InterPro; IPR007087; Znf_C2H2.
DR   Pfam; PF00096; zf-C2H2; 1.
DR   SMART; SM00355; ZnF_C2H2; 2.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
KW   Complete proteome; Developmental protein; Hypothetical protein;
KW   Metal-binding; Nuclear protein; Zinc; Zinc-finger.
SQ   SEQUENCE   355 AA;  39236 MW;  56B98564B2926062 CRC64;

  Query Match           33.1%;  Score 58.5;  DB 2;  Length 355;
  Best Local Similarity  46.2%;  Pred. No. 30;
  Matches   18;  Conservative    5;  Mismatches    9;  Indels    7;  Gaps    2;

Qy          3 SSLFSPSSSS----YSSLFTAK---PTRLLSPKPKFTFS 34
              ||:|: || |    ||:: ||    || | ||| :|  |
Db         73 SSMFAASSPSAAATYSTVTTAALVVPTTLQSPKREFVCS 111


RESULT 14
Q75WN1_CHAOB
ID   Q75WN1_CHAOB PRELIMINARY;      PRT;   605 AA.
AC   Q75WN1;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Limonene/borneol synthase.
GN   Name=co1;
OS   Chamaecyparis obtusa (Japanese cypress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Cupressaceae;
OC   Chamaecyparis.
OX   NCBI_TaxID=13415;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Needle;
RA   Katoh S., Hyatt D., Ogawa K., Fujii M., Furuno T., Croteau R.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB120957; BAC92722.1; -; mRNA.
DR   GO; GO:0016829; F:lyase activity; IEA.
DR   GO; GO:0000287; F:magnesium ion binding; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR005630; Terpene_synth_C.
DR   InterPro; IPR001906; Terp_synth-like.
DR   Pfam; PF01397; Terpene_synth; 1.
DR   Pfam; PF03936; Terpene_synth_C; 1.
SQ   SEQUENCE   605 AA;  70309 MW;  CE4DD12D6A4462F2 CRC64;

  Query Match           32.8%;  Score 58;  DB 2;  Length 605;
  Best Local Similarity  44.4%;  Pred. No. 64;
  Matches   12;  Conservative    5;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          4 SLFSPSSSSYSSLFTAKPTRLLSPKPK 30
              :| | |||::::    :||| |  | ||
Db          2 ALISLSSSAFTFCLKSKPTHLSKPSPK 28


RESULT 15
Q86H13_DICDI
ID   Q86H13_DICDI PRELIMINARY;      PRT;   771 AA.
AC   Q86H13;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Putative cyclic nucleotide phosphodiesterase (Hypothetical protein
DE   Pde4).
GN   Name=Pde4; ORFNames=DDB0216197;
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Bader S., Van Haastert P.J.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AX4;
RA   Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA   Sucgang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA   Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA   Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA   Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA   Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA   Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA   Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA   Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA   Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA   Wardroper A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA   Loulseged H., Mungall K., Oliver K., Price C., Quail M.A.,
RA   Urushihara H., Hernandez J., Rabbinowitsch E., Steffen D., Sanders M.,
RA   Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA   Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA   Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA   Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
RA   Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.;
RT   "The genome of the social amoeba Dictyostelium discoideum.";
RL   Nature 0:0-0(2005).
DR   EMBL; AY211984; AAO59486.1; -; mRNA.
DR   EMBL; AAFI01000196; EAL62868.1; -; Genomic_DNA.
DR   HSSP; Q08499; 1OYN.
DR   DictyBase; DDB0216197; Pde4.
DR   GO; GO:0004114; F:3',5'-cyclic-nucleotide phosphodiesterase a. . .; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR003607; Met_phos_hydro.
DR   InterPro; IPR002073; PDEase.
DR   Pfam; PF00233; PDEase_I; 1.
DR   PRINTS; PR00387; PDIESTERASE1.
DR   SMART; SM00471; HDc; 1.
DR   PROSITE; PS00126; PDEASE_I; UNKNOWN_1.
KW   Hypothetical protein.
SQ   SEQUENCE   771 AA;  88908 MW;  2036F7D8DC9D6E07 CRC64;

  Query Match           32.8%;  Score 58;  DB 2;  Length 771;
```

```
  Best Local Similarity   51.9%;  Pred. No. 85;
  Matches   14;  Conservative    2;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          8 PSSSSYSSLFTAKPTRLLSPKPKFTFS 34
              ||||| ||  |  ||  ::| |  | ||
Db        479 PSSSSSSSSSTTIPTSTITPTPNSTTS 505


Search completed: March  7, 2006, 21:58:12
Job time : 181.2 secs
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2006, 21:59:17 ; Search time 140.4 Seconds
(without alignments)
107.136 Million cell updates/sec

Title: US-10-751-235-11
Perfect score: 177
Sequence: 1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	177	100.0	36	5	US-10-751-235-11	Sequence 11, Appl
2	177	100.0	539	5	US-10-751-235-4	Sequence 4, Appli
3	177	100.0	560	5	US-10-739-930-6317	Sequence 6317, Ap
4	63	35.6	260	4	US-10-310-154-688	Sequence 688, App
5	63	35.6	260	5	US-10-732-923-601	Sequence 601, App
6	63	35.6	262	5	US-10-732-923-21136	Sequence 21136, A
7	62.5	35.3	2016	6	US-11-097-143-18525	Sequence 18525, A
8	61.5	34.7	309	5	US-10-732-923-5867	Sequence 5867, Ap
9	61.5	34.7	424	5	US-10-732-923-5866	Sequence 5866, Ap
10	61	34.5	82	4	US-10-425-115-348345	Sequence 348345,
11	61	34.5	389	4	US-10-424-599-230527	Sequence 230527,
12	59.5	33.6	319	3	US-09-870-406A-57	Sequence 57, Appl
13	59.5	33.6	319	4	US-10-159-901-57	Sequence 57, Appl
14	59.5	33.6	319	5	US-10-739-930-5731	Sequence 5731, Ap
15	59.5	33.6	525	4	US-10-424-599-260318	Sequence 260318,
16	59	33.3	1545	6	US-11-097-143-23523	Sequence 23523, A
17	58.5	33.1	78	4	US-10-424-599-278783	Sequence 278783,
18	58	32.8	109	4	US-10-425-115-257160	Sequence 257160,
19	58	32.8	131	4	US-10-425-115-257159	Sequence 257159,
20	58	32.8	160	4	US-10-425-115-257153	Sequence 257153,
21	57.5	32.5	215	4	US-10-424-599-201457	Sequence 201457,
22	57.5	32.5	283	4	US-10-424-599-261750	Sequence 261750,
23	56.5	31.9	131	4	US-10-424-599-236698	Sequence 236698,
24	56	31.6	48	5	US-10-450-763-35222	Sequence 35222, A
25	56	31.6	204	4	US-10-424-599-196060	Sequence 196060,
26	56	31.6	883	6	US-11-097-143-13482	Sequence 13482, A
27	55.5	31.4	370	4	US-10-424-599-245828	Sequence 245828,
28	55.5	31.4	370	4	US-10-425-114-37736	Sequence 37736, A
29	55.5	31.4	706	4	US-10-424-599-218604	Sequence 218604,
30	55	31.1	97	4	US-10-424-599-196064	Sequence 196064,
31	55	31.1	116	4	US-10-424-599-159373	Sequence 159373,
32	55	31.1	246	4	US-10-424-599-188070	Sequence 188070,
33	55	31.1	254	4	US-10-424-599-191224	Sequence 191224,
34	55	31.1	282	4	US-10-424-599-147306	Sequence 147306,
35	55	31.1	741	4	US-10-424-599-158653	Sequence 158653,
36	55	31.1	1536	5	US-10-498-428-2	Sequence 2, Appli
37	54.5	30.8	416	4	US-10-374-780A-1314	Sequence 1314, Ap
38	54.5	30.8	416	4	US-10-412-699B-1459	Sequence 1459, Ap
39	54.5	30.8	527	4	US-10-424-599-194264	Sequence 194264,
40	54.5	30.8	2027	4	US-10-234-961-2	Sequence 2, Appli
41	54	30.5	87	4	US-10-424-599-256870	Sequence 256870,
42	54	30.5	167	4	US-10-424-599-224645	Sequence 224645,
43	54	30.5	170	5	US-10-732-923-21141	Sequence 21141, A
44	54	30.5	195	4	US-10-424-599-265695	Sequence 265695,
45	54	30.5	205	4	US-10-094-458A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-10-751-235-11
; Sequence 11, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: DellaPenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; TITLE OF INVENTION: Metabolism in Plants
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-751-235-11

Query Match 100.0%; Score 177; DB 5; Length 36;
Best Local Similarity 100.0%; Pred. No. 1e-15;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36
||||||||||||||||||||||||||||||||||||
Db 1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36

RESULT 2
US-10-751-235-4
; Sequence 4, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: DellaPenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; TITLE OF INVENTION: Metabolism in Plants
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 539
; TYPE: PRT

```
;   ORGANISM: Arabidopsis thaliana
US-10-751-235-4

  Query Match             100.0%;  Score 177;  DB 5;  Length 539;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-14;
  Matches   36;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36
              |||||||||||||||||||||||||||||||||||||
Db          1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36


RESULT 3
US-10-739-930-6317
; Sequence 6317, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
;  APPLICANT: Kovalic, David K.
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
;  TITLE OF INVENTION:  PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
;  FILE REFERENCE: 38-21(53377)B
;  CURRENT APPLICATION NUMBER: US/10/739,930
;  CURRENT FILING DATE:  2003-12-18
;  NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 6317
;   LENGTH: 560
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   OTHER INFORMATION: Clone ID: ARATH-23APR03-C3001_1.p
US-10-739-930-6317

  Query Match             100.0%;  Score 177;  DB 5;  Length 560;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-14;
  Matches   36;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 36
              |||||||||||||||||||||||||||||||||||||
Db         22 MESSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFSIR 57


RESULT 4
US-10-310-154-688
; Sequence 688, Application US/10310154
; Publication No. US20030233670A1
; GENERAL INFORMATION:
;  APPLICANT: Edgerton, Michael D
;  APPLICANT:  Chomet, Paul S.
;  APPLICANT:  Adams, Thomas H
;  APPLICANT:  Ruff, Thomas G.
;  APPLICANT:  Agarwal, Ameeta K.
;  APPLICANT:  Ahrens, Jeffrey E.
;  APPLICANT:  Ball, James A.
;  APPLICANT:  Banu, G.
;  APPLICANT:  Bell, Erin
;  APPLICANT:  Boddupalli, Raghava
;  APPLICANT:  Deikman, Jill
;  APPLICANT:  Deng, Molian
;  APPLICANT:  Dong, Jinzhuo
;  APPLICANT:  Duff, Stephen M.
;  APPLICANT:  Galligan, Meghan M.
;  APPLICANT:  Hinchey, Brenda S.
;  APPLICANT:  Huang, Shihshieh
;  APPLICANT:  Johnson, G. Richard
;  APPLICANT:  Jung, Vincent
;  APPLICANT:  Kretzmer, Keith A
;  APPLICANT:  Laccetti, Lucille B.
;  APPLICANT:  Lai, Chao-Qiang
;  APPLICANT:  Lee, Gary
;  APPLICANT:  Lin, Jie-Yi
;  APPLICANT:  Liu, Jingdong
;  APPLICANT:  Lu, Bin
;  APPLICANT:  Luethy, Michael M.
;  APPLICANT:  Lund, Adrian
;  APPLICANT:  Madson, Linda L.
;  APPLICANT:  Malloy, Kathleen A.
;  APPLICANT:  McKiel, Christine L.
;  APPLICANT:  Miller, Philip W.
;  APPLICANT:  Padmavathi, Manchikanti
;  APPLICANT:  Parnell, Laurence D.
;  APPLICANT:  Start, William G.
;  APPLICANT:  Tennesen, Dan
;  APPLICANT:  Vidya, K.R.
;  APPLICANT:  Wang, Haiyun
;  APPLICANT:  Xin, Zhanguo
;  APPLICANT:  Xu, Nanfei
;  APPLICANT:  Yang, Chunzhi
;  APPLICANT:  Zeng, Xiaoping
;  APPLICANT:  Zhang, Qiang
;  APPLICANT:  Zhao, Yajuan
;  APPLICANT:  Zhou, Li
;  TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
;  FILE REFERENCE: 38-15(52796)B
;  CURRENT APPLICATION NUMBER: US/10/310,154
;  CURRENT FILING DATE:  2002-12-04
;  PRIOR APPLICATION NUMBER: 60/337,358
;  PRIOR FILING DATE: 2001-12-04
;  NUMBER OF SEQ ID NOS: 736
; SEQ ID NO 688
;   LENGTH: 260
;   TYPE: PRT
;   ORGANISM: Glycine max
US-10-310-154-688

  Query Match             35.6%;  Score 63;  DB 4;  Length 260;
  Best Local Similarity   46.2%;  Pred. No. 5.7;
  Matches   12;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         10 SSSYSSLFTAKPTRLLSPKPKFTFSI 35
              |:: :|||:| || | ||||  :  :
Db          4 SATSASLFSANPTPLFSPKPSLSLHL 29


RESULT 5
US-10-732-923-601
; Sequence 601, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
;  APPLICANT: Edgerton, Michael D
;  TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
;  FILE REFERENCE: 38-15(52796)C
;  CURRENT APPLICATION NUMBER: US/10/732,923
;  CURRENT FILING DATE:  2003-12-10
;  PRIOR APPLICATION NUMBER: 10/310,154
;  PRIOR FILING DATE: 2002-12-04
;  NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 601
;   LENGTH: 260
;   TYPE: PRT
;   ORGANISM: Glycine max
US-10-732-923-601

  Query Match             35.6%;  Score 63;  DB 5;  Length 260;
  Best Local Similarity   46.2%;  Pred. No. 5.7;
  Matches   12;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         10 SSSYSSLFTAKPTRLLSPKPKFTFSI 35
              |:: :|||:| || | ||||  :  :
Db          4 SATSASLFSANPTPLFSPKPSLSLHL 29


RESULT 6
US-10-732-923-21136
; Sequence 21136, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Edgerton, Michael D
;  TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
;  FILE REFERENCE: 38-15(52796)C
;  CURRENT APPLICATION NUMBER: US/10/732,923
;  CURRENT FILING DATE:  2003-12-10
;  PRIOR APPLICATION NUMBER: 10/310,154
;  PRIOR FILING DATE: 2002-12-04
;  NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 21136
;   LENGTH: 262
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(262)
;   OTHER INFORMATION: unsure at all Xaa locations
US-10-732-923-21136

  Query Match           35.6%;  Score 63;  DB 5;  Length 262;
  Best Local Similarity  46.2%;  Pred. No. 5.8;
  Matches   12;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         10 SSSYSSLFTAKPTRLLSPKPKFTFSI 35
              |:: :|||:| || | |||| :  :
Db          4 SATSASLFSANPTPLFSPKPSLSLHL 29


RESULT 7
US-11-097-143-18525
; Sequence 18525, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
;  APPLICANT: Venter, J. Craig
;  APPLICANT:  et al.
;  TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
;  TITLE OF INVENTION:  ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
;  TITLE OF INVENTION:  DROSOPHILA GENES.
;  FILE REFERENCE: CL000728
;  CURRENT APPLICATION NUMBER: US/11/097,143
;  CURRENT FILING DATE:  2005-04-04
;  PRIOR APPLICATION NUMBER: 60/157,832
;  PRIOR FILING DATE: 1999-10-05
;  PRIOR APPLICATION NUMBER: 60/160,191
;  PRIOR FILING DATE: 1999-10-19
;  PRIOR APPLICATION NUMBER: 60/161,932
;  PRIOR FILING DATE: 1999-10-28
;  PRIOR APPLICATION NUMBER: 60/164,769
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/173,383
;  PRIOR FILING DATE: 1999-12-28
;  PRIOR APPLICATION NUMBER: 60/175,693
;  PRIOR FILING DATE: 2000-01-12
;  PRIOR APPLICATION NUMBER: 60/184,831
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/191,637
;  PRIOR FILING DATE: 2000-03-23
;  NUMBER OF SEQ ID NOS: 43008
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18525
;   LENGTH: 2016
;   TYPE: PRT
;   ORGANISM: DROSOPHILA
US-11-097-143-18525

  Query Match           35.3%;  Score 62.5;  DB 6;  Length 2016;
  Best Local Similarity  47.4%;  Pred. No. 65;
  Matches   18;  Conservative    4;  Mismatches   11;  Indels    5;  Gaps    2;

Qy          3 SSLFSPSS-SSYSSLFTAKP----TRLLSPKPKFTFSI 35
              |||||||   | ||||   |    | :|:| |  | ::
Db        986 SSLFSPSQVDSTSSLFAPPPTSANTTILTPAPDTTSTL 1023


RESULT 8
US-10-732-923-5867
; Sequence 5867, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
;  APPLICANT: Edgerton, Michael D
;  TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
;  FILE REFERENCE: 38-15(52796)C
;  CURRENT APPLICATION NUMBER: US/10/732,923
;  CURRENT FILING DATE:  2003-12-10
;  PRIOR APPLICATION NUMBER: 10/310,154
;  PRIOR FILING DATE: 2002-12-04
;  NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 5867
;   LENGTH: 309
;   TYPE: PRT
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(309)
;   OTHER INFORMATION: unsure at all Xaa locations
US-10-732-923-5867

  Query Match           34.7%;  Score 61.5;  DB 5;  Length 309;
  Best Local Similarity  43.8%;  Pred. No. 11;
  Matches   14;  Conservative    6;  Mismatches   11;  Indels    1;  Gaps    1;

Qy          2 ESSLFSPSSSSYSSLFTAKPTRLLSP-KPKFT 32
              :| : :||||||:| | | |    :|  | |:
Db         31 DSYMATPSSSSYASSFAASPLPTTAPASPSFS 62


RESULT 9
US-10-732-923-5866
; Sequence 5866, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
;  APPLICANT: Edgerton, Michael D
;  TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
;  FILE REFERENCE: 38-15(52796)C
;  CURRENT APPLICATION NUMBER: US/10/732,923
;  CURRENT FILING DATE:  2003-12-10
;  PRIOR APPLICATION NUMBER: 10/310,154
;  PRIOR FILING DATE: 2002-12-04
;  NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 5866
;   LENGTH: 424
;   TYPE: PRT
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(424)
;   OTHER INFORMATION: unsure at all Xaa locations
US-10-732-923-5866

  Query Match           34.7%;  Score 61.5;  DB 5;  Length 424;
  Best Local Similarity  43.8%;  Pred. No. 15;
  Matches   14;  Conservative    6;  Mismatches   11;  Indels    1;  Gaps    1;

Qy          2 ESSLFSPSSSSYSSLFTAKPTRLLSP-KPKFT 32
              :| : :||||||:| | | |    :|  | |:
Db        150 DSYMATPSSSSYASSFAASPLPTTAPASPSFS 181


RESULT 10
US-10-425-115-348345
; Sequence 348345, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
```

```
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 348345
;   LENGTH: 82
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_80854C.1.pep
US-10-425-115-348345

  Query Match             34.5%;  Score 61;  DB 4;  Length 82;
  Best Local Similarity   43.5%;  Pred. No. 2.9;
  Matches   10;  Conservative    6;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         13 YSSLFTAKPTRLLSPKPKFTFSI 35
              :  :|:  ||::| | |: ||||
Db         59 HKKIFSLNPTKILIPTPRVTFSI 81


RESULT 11
US-10-424-599-230527
; Sequence 230527, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 230527
;   LENGTH: 389
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_50188C.1.pep
US-10-424-599-230527

  Query Match             34.5%;  Score 61;  DB 4;  Length 389;
  Best Local Similarity   46.9%;  Pred. No. 16;
  Matches   15;  Conservative    4;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          3 SSLFSPSSSSYSSLFTAKPTRLLSPKPKFTFS 34
              ||  |||::: ||   : | ||  |||   ||
Db         50 SSCSSPSAAASSSSSPSSPFRLRLPKPPTAFS 81


RESULT 12
US-09-870-406A-57
; Sequence 57, Application US/09870406A
; Publication No. US20030104379A1
; GENERAL INFORMATION:
;  APPLICANT: LAGARIAS, JOHN
;  APPLICANT:  KOICHI, TAKAYUKI
;  APPLICANT:  FRANKENBERG, NICOLE
;  APPLICANT:  GAMBETTA, GREGORY
;  APPLICANT:  MONTGOMERY, BERONDA
;  TITLE OF INVENTION: HY2 FAMILY OF BILIN REDUCTASES
;  FILE REFERENCE: 407T-907720US
;  CURRENT APPLICATION NUMBER: US/09/870,406A
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: 60/271,758
;  PRIOR FILING DATE: 2001-02-26
;  PRIOR APPLICATION NUMBER: 60/210,286
;  PRIOR FILING DATE: 2000-06-08
;  NUMBER OF SEQ ID NOS: 57
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 57
;   LENGTH: 319
;   TYPE: PRT
;   ORGANISM: Arapidopsis thaliana
US-09-870-406A-57

  Query Match             33.6%;  Score 59.5;  DB 3;  Length 319;
  Best Local Similarity   43.8%;  Pred. No. 20;
  Matches   14;  Conservative    7;  Mismatches    8;  Indels    3;  Gaps    1;

Qy          3 SSLFSPSSSSYSSLFTAKPTRL---LSPKPKF 31
              ::|:| || || |  |:||:|    | |: :|
Db          7 NTLYSSSSPSYLSPLTSKPSRFSKNLRPRAQF 38


RESULT 13
US-10-159-901-57
; Sequence 57, Application US/10159901
; Publication No. US20030073235A1
; GENERAL INFORMATION:
;  APPLICANT: LAGARIAS, JOHN
;  APPLICANT:  KOICHI, TAKAYUKI
;  APPLICANT:  FRANKENBERG, NICOLE
;  APPLICANT:  GAMBETTA, GREGORY
;  APPLICANT:  MONTGOMERY, BERONDA
;  TITLE OF INVENTION: LIGHT CONTROLLED GENE EXPRESSION UTILIZING HETEROLOGOUS PHYTOCHRON
;  FILE REFERENCE: 407T-907731US
;  CURRENT APPLICATION NUMBER: US/10/159,901
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: 60/294,463
;  PRIOR FILING DATE: 2001-05-29
;  NUMBER OF SEQ ID NOS: 57
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 57
;   LENGTH: 319
;   TYPE: PRT
;   ORGANISM: Arapidopsis thaliana
US-10-159-901-57

  Query Match             33.6%;  Score 59.5;  DB 4;  Length 319;
  Best Local Similarity   43.8%;  Pred. No. 20;
  Matches   14;  Conservative    7;  Mismatches    8;  Indels    3;  Gaps    1;

Qy          3 SSLFSPSSSSYSSLFTAKPTRL---LSPKPKF 31
              ::|:| || || |  |:||:|    | |: :|
Db          7 NTLYSSSSPSYLSPLTSKPSRFSKNLRPRAQF 38


RESULT 14
US-10-739-930-5731
; Sequence 5731, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
;  APPLICANT: Kovalic, David K.
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
;  TITLE OF INVENTION:  PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
;  FILE REFERENCE: 38-21(53377)B
;  CURRENT APPLICATION NUMBER: US/10/739,930
;  CURRENT FILING DATE:  2003-12-18
;  NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 5731
;   LENGTH: 319
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   OTHER INFORMATION: Clone ID: ARATH-23APR03-C1272_1.p
US-10-739-930-5731

  Query Match             33.6%;  Score 59.5;  DB 5;  Length 319;
  Best Local Similarity   43.8%;  Pred. No. 20;
  Matches   14;  Conservative    7;  Mismatches    8;  Indels    3;  Gaps    1;
```

```
Qy          3 SSLFSPSSSSYSSLFTAKPTRL---LSPKPKF 31
              ::|:| || || |  |:||:|    | |: :|
Db          7 NTLYSSSSPSYLSPLTSKPSRFSKNLRPRAQF 38


RESULT 15
US-10-424-599-260318
; Sequence 260318, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 260318
;   LENGTH: 525
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(525)
;   OTHER INFORMATION: unsure at all Xaa locations
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_77090C.1.pep
US-10-424-599-260318

  Query Match             33.6%;  Score 59.5;  DB 4;  Length 525;
  Best Local Similarity   40.0%;  Pred. No. 36;
  Matches   14;  Conservative    9;  Mismatches    7;  Indels    5;  Gaps    2;

Qy          5 LFSPSSSSYSSLF--TAKPTRLL---SPKPKFTFS 34
              :::| | |::|:   : |  |||   :||||| |:
Db         49 IYTPKSPSFTSVLDSSGKNQRLLVPSTPKPKFIFT 83


Search completed: March  7, 2006, 22:04:14
Job time : 140.4 secs
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2006, 11:24:53 ; Search time 165 Seconds
(without alignments)
25.323 Million cell updates/sec

Title: US-10-751-235-14
Perfect score: 33
Sequence: 1 FXXGXXXCXG 10

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	100.0	10	3	US-09-739-254-129	Sequence 129, App
2	33	100.0	10	3	US-09-904-615-129	Sequence 129, App
3	33	100.0	10	4	US-10-067-534-15	Sequence 15, Appl
4	33	100.0	10	4	US-10-210-965-5	Sequence 5, Appli
5	33	100.0	10	4	US-10-054-988-129	Sequence 129, App
6	33	100.0	10	4	US-10-236-433-11	Sequence 11, Appl
7	33	100.0	10	5	US-10-477-526-7	Sequence 7, Appli
8	33	100.0	10	5	US-10-477-526-8	Sequence 8, Appli
9	33	100.0	10	5	US-10-751-235-14	Sequence 14, Appl
10	33	100.0	10	5	US-10-751-235-15	Sequence 15, Appl
11	33	100.0	11	4	US-10-097-559-28	Sequence 28, Appl
12	33	100.0	11	5	US-10-804-772-26	Sequence 26, Appl
13	33	100.0	11	5	US-10-804-772-29	Sequence 29, Appl
14	33	100.0	14	4	US-10-373-877-10	Sequence 10, Appl
15	33	100.0	14	4	US-10-373-877-19	Sequence 19, Appl
16	33	100.0	14	4	US-10-373-877-20	Sequence 20, Appl
17	33	100.0	14	4	US-10-373-877-21	Sequence 21, Appl
18	33	100.0	14	4	US-10-373-877-22	Sequence 22, Appl
19	33	100.0	14	4	US-10-373-877-23	Sequence 23, Appl
20	33	100.0	14	4	US-10-373-877-24	Sequence 24, Appl
21	33	100.0	14	4	US-10-373-877-25	Sequence 25, Appl
22	33	100.0	14	5	US-10-865-478-774	Sequence 774, App
23	33	100.0	15	3	US-09-957-674-13	Sequence 13, Appl
24	33	100.0	15	4	US-10-257-378-1	Sequence 1, Appli
25	33	100.0	15	5	US-10-882-241-17	Sequence 17, Appl
26	33	100.0	16	4	US-10-163-198-90	Sequence 90, Appl
27	33	100.0	16	5	US-10-926-683-1257	Sequence 1257, Ap
28	33	100.0	18	2	US-08-647-444-12	Sequence 12, Appl
29	33	100.0	20	5	US-10-855-595-6	Sequence 6, Appli
30	33	100.0	20	5	US-10-855-595-7	Sequence 7, Appli
31	33	100.0	20	5	US-10-855-595-8	Sequence 8, Appli
32	33	100.0	20	5	US-10-855-595-9	Sequence 9, Appli
33	33	100.0	20	5	US-10-855-595-10	Sequence 10, Appl
34	33	100.0	20	5	US-10-855-532-6	Sequence 6, Appli
35	33	100.0	20	5	US-10-855-532-7	Sequence 7, Appli
36	33	100.0	20	5	US-10-855-532-8	Sequence 8, Appli
37	33	100.0	20	5	US-10-855-532-9	Sequence 9, Appli
38	33	100.0	20	5	US-10-855-532-10	Sequence 10, Appl
39	33	100.0	21	4	US-10-021-425-25	Sequence 25, Appl
40	33	100.0	21	5	US-10-900-856-28	Sequence 28, Appl
41	33	100.0	23	4	US-10-103-196-31	Sequence 31, Appl
42	33	100.0	30	3	US-09-749-637A-126	Sequence 126, App
43	33	100.0	30	3	US-09-749-637A-129	Sequence 129, App
44	33	100.0	30	3	US-09-749-637A-283	Sequence 283, App
45	33	100.0	30	4	US-10-433-485A-1	Sequence 1, Appli
46	33	100.0	30	5	US-10-839-227-126	Sequence 126, App
47	33	100.0	30	5	US-10-839-227-129	Sequence 129, App
48	33	100.0	30	5	US-10-839-227-283	Sequence 283, App
49	33	100.0	32	4	US-10-437-963-159510	Sequence 159510,
50	33	100.0	35	4	US-10-351-641-487	Sequence 487, App
51	33	100.0	35	4	US-10-351-641-488	Sequence 488, App
52	33	100.0	35	4	US-10-351-641-489	Sequence 489, App
53	33	100.0	35	4	US-10-351-641-490	Sequence 490, App
54	33	100.0	35	4	US-10-351-641-491	Sequence 491, App
55	33	100.0	35	4	US-10-351-641-492	Sequence 492, App
56	33	100.0	35	4	US-10-351-641-493	Sequence 493, App
57	33	100.0	35	4	US-10-351-641-523	Sequence 523, App
58	33	100.0	35	4	US-10-351-641-524	Sequence 524, App
59	33	100.0	35	4	US-10-351-641-525	Sequence 525, App
60	33	100.0	35	4	US-10-351-641-526	Sequence 526, App
61	33	100.0	35	4	US-10-351-641-527	Sequence 527, App
62	33	100.0	35	4	US-10-351-641-528	Sequence 528, App
63	33	100.0	35	4	US-10-351-641-529	Sequence 529, App
64	33	100.0	39	4	US-10-147-140-54	Sequence 54, Appl
65	33	100.0	40	3	US-09-894-882-348	Sequence 348, App
66	33	100.0	40	3	US-09-894-882-482	Sequence 482, App
67	33	100.0	40	5	US-10-894-314A-348	Sequence 348, App
68	33	100.0	40	5	US-10-894-314A-482	Sequence 482, App
69	33	100.0	41	4	US-10-425-115-244467	Sequence 244467,
70	33	100.0	42	4	US-10-424-599-143366	Sequence 143366,
71	33	100.0	42	4	US-10-424-599-256014	Sequence 256014,
72	33	100.0	42	4	US-10-425-115-238482	Sequence 238482,
73	33	100.0	43	4	US-10-424-599-187772	Sequence 187772,
74	33	100.0	44	4	US-10-424-599-271553	Sequence 271553,
75	33	100.0	44	4	US-10-425-115-239272	Sequence 239272,
76	33	100.0	46	4	US-10-267-682-208	Sequence 208, App
77	33	100.0	46	4	US-10-267-748-208	Sequence 208, App
78	33	100.0	49	4	US-10-424-599-157117	Sequence 157117,
79	33	100.0	49	4	US-10-425-115-238518	Sequence 238518,
80	33	100.0	49	4	US-10-425-115-240869	Sequence 240869,
81	33	100.0	49	4	US-10-425-115-324478	Sequence 324478,
82	33	100.0	49	4	US-10-425-115-351931	Sequence 351931,
83	33	100.0	50	4	US-10-425-115-260285	Sequence 260285,
84	33	100.0	51	4	US-10-424-599-167757	Sequence 167757,
85	33	100.0	51	4	US-10-424-599-177489	Sequence 177489,
86	33	100.0	51	4	US-10-424-599-214094	Sequence 214094,
87	33	100.0	52	3	US-09-864-761-45740	Sequence 45740, A
88	33	100.0	52	4	US-10-424-599-254214	Sequence 254214,
89	33	100.0	52	5	US-10-926-683-1061	Sequence 1061, Ap
90	33	100.0	54	4	US-10-424-599-273991	Sequence 273991,
91	33	100.0	54	4	US-10-437-963-185080	Sequence 185080,
92	33	100.0	55	3	US-09-864-761-34668	Sequence 34668, A
93	33	100.0	55	4	US-10-106-698-6691	Sequence 6691, Ap
94	33	100.0	55	4	US-10-425-115-250988	Sequence 250988,
95	33	100.0	56	4	US-10-425-115-361393	Sequence 361393,
96	33	100.0	57	4	US-10-424-599-199843	Sequence 199843,
97	33	100.0	57	4	US-10-425-115-261158	Sequence 261158,
98	33	100.0	59	4	US-10-424-599-176580	Sequence 176580,
99	33	100.0	59	4	US-10-425-115-280265	Sequence 280265,
100	33	100.0	59	4	US-10-425-115-314843	Sequence 314843,

ALIGNMENTS

```
RESULT 1
US-09-739-254-129
; Sequence 129, Application US/09739254
; Patent No. US20010021700A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 49 Human Secreted Proteins
;  FILE REFERENCE: PZ032P1
;  CURRENT APPLICATION NUMBER: US/09/739,254
;  CURRENT FILING DATE: 2000-12-19
;  EARLIER APPLICATION NUMBER: 09/511,554
;  EARLIER FILING DATE: 2000-02-23
;  EARLIER APPLICATION NUMBER: PCT/US99/19330
;  EARLIER FILING DATE: 1999-08-24
;  EARLIER APPLICATION NUMBER: 60/097,917
;  EARLIER FILING DATE: 1998-08-25
;  EARLIER APPLICATION NUMBER: 60/098,634
;  EARLIER FILING DATE: 1998-08-31
;  NUMBER OF SEQ ID NOS: 170
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 129
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-739-254-129

  Query Match             100.0%;  Score 33;  DB 3;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 3.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FSLGRRHCLG 10


RESULT 2
US-09-904-615-129
; Sequence 129, Application US/09904615
; Patent No. US20020026040A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 49 Human Secreted Proteins
;  FILE REFERENCE: PZ032P1
;  CURRENT APPLICATION NUMBER: US/09/904,615
;  CURRENT FILING DATE:  2001-07-16
;  PRIOR APPLICATION NUMBER: 09/511,554
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/097,917
;  PRIOR FILING DATE: 1998-08-25
;  PRIOR APPLICATION NUMBER: 60/098,634
;  PRIOR FILING DATE: 1998-08-31
;  NUMBER OF SEQ ID NOS: 170
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 129
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-904-615-129

  Query Match             100.0%;  Score 33;  DB 3;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 3.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FSLGRRHCLG 10


RESULT 3
US-10-067-534-15
; Sequence 15, Application US/10067534
; Publication No. US20020187538A1
; GENERAL INFORMATION:
;  APPLICANT: Essenberg, Margaret K.
;  APPLICANT:  Chen, Xiao-Ya
;  APPLICANT:  Luo, Ping
;  APPLICANT:  Wang, Yan-Hong
;  TITLE OF INVENTION: cDNA Clone of  (+)-Delta-Cadinene-8-Hydroxylase Gene from Cotton P
;  FILE REFERENCE: 006602-113
;  CURRENT APPLICATION NUMBER: US/10/067,534
;  CURRENT FILING DATE:  2002-02-07
;  PRIOR APPLICATION NUMBER: US 60/267,160
;  PRIOR FILING DATE: 2001-02-07
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Gossypium arboreum
US-10-067-534-15

  Query Match             100.0%;  Score 33;  DB 4;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 3.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGSGRRMCAG 10


RESULT 4
US-10-210-965-5
; Sequence 5, Application US/10210965
; Publication No. US20030078404A1
; GENERAL INFORMATION:
;  APPLICANT: Glucksmann, Maria A.
;  TITLE OF INVENTION: 33297, A Human Cytochrome P450 Family
;  TITLE OF INVENTION:  Member and Uses Therefor
;  FILE REFERENCE: MPI01-091P1RM
;  CURRENT APPLICATION NUMBER: US/10/210,965
;  CURRENT FILING DATE:  2002-08-02
;  PRIOR APPLICATION NUMBER: 60/312428
;  PRIOR FILING DATE: 2001-08-15
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: consensus pattern, Prosite Accession No. US20030078404A1 PS00086
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 1
;   OTHER INFORMATION: Phe can be Trp.
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 2
;   OTHER INFORMATION: Ser can be Gly, Asn, or His.
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 3
;   OTHER INFORMATION: Xaa can be any amino acid.
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 4
;   OTHER INFORMATION: Gly can be Asp.
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 5
;   OTHER INFORMATION: Xaa can be any amino acid.
```

```
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (6)...(6)
;   OTHER INFORMATION: Arg can be Lys, His, Pro, or Thr.
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (7)...(7)
;   OTHER INFORMATION: Xaa can be any amino acid.
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (8)...(8)
;   OTHER INFORMATION: Cys is involved in binding heme iron
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (9)...(9)
;   OTHER INFORMATION: Leu can be Ile, Val, Met, Phe, Ala, or Pro.
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (10)...(10)
;   OTHER INFORMATION: Gly can be Ala or Asp.
US-10-210-965-5

  Query Match           100.0%;  Score 33;  DB 4;  Length 10;
  Best Local Similarity  70.0%;  Pred. No. 3.5e+02;
  Matches    7;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |:|||:||:|
Db          1 FSXGXRXCLG 10


RESULT 5
US-10-054-988-129
; Sequence 129, Application US/10054988
; Publication No. US20030087341A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 49 Human Secreted Proteins
;  FILE REFERENCE: PZ032P1
;  CURRENT APPLICATION NUMBER: US/10/054,988
;  CURRENT FILING DATE:  2002-01-25
;  PRIOR APPLICATION NUMBER: 09/904,615
;  PRIOR FILING DATE: 2001-07-16
;  PRIOR APPLICATION NUMBER: 09/511,554
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/097,917
;  PRIOR FILING DATE: 1998-08-25
;  PRIOR APPLICATION NUMBER: 60/098,634
;  PRIOR FILING DATE: 1998-08-31
;  NUMBER OF SEQ ID NOS: 170
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 129
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-054-988-129

  Query Match           100.0%;  Score 33;  DB 4;  Length 10;
  Best Local Similarity  40.0%;  Pred. No. 3.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FSLGRRHCLG 10


RESULT 6
US-10-236-433-11
; Sequence 11, Application US/10236433
; Publication No. US20030100025A1
; GENERAL INFORMATION:
;  APPLICANT: O'Connor, Michael B.
;  APPLICANT:  Gilbert, Lawrence I.
;  APPLICANT:  Warren, James T.
;  TITLE OF INVENTION: Insecticide Targets
;  FILE REFERENCE: 09531-070001
;  CURRENT APPLICATION NUMBER: US/10/236,433
;  CURRENT FILING DATE:  2002-12-10
;  PRIOR APPLICATION NUMBER: 60/318,006
;  PRIOR FILING DATE: 2001-09-07
;  PRIOR APPLICATION NUMBER: 60/317,890
;  PRIOR FILING DATE: 2001-09-07
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: P450 motif
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 2,3,5,6,7,9
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-10-236-433-11

  Query Match           100.0%;  Score 33;  DB 4;  Length 10;
  Best Local Similarity  100.0%;  Pred. No. 3.5e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              ||||||||||
Db          1 FXXGXXXCXG 10


RESULT 7
US-10-477-526-7
; Sequence 7, Application US/10477526
; Publication No. US20050009021A1
; GENERAL INFORMATION:
;  APPLICANT: Wisniewski, Jan
;  APPLICANT:  Ramshaw, Heather
;  APPLICANT:  Petkovich, Martin P.
;  TITLE OF INVENTION: A Retinoic Acid Metabolizing Cytochrome P450
;  FILE REFERENCE: 11812-91
;  CURRENT APPLICATION NUMBER: US/10/477,526
;  CURRENT FILING DATE:  2003-11-21
;  PRIOR APPLICATION NUMBER: US 60/292,531
;  PRIOR FILING DATE: 2001-05-24
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (2)..(3)
;   OTHER INFORMATION: x=any residue
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (9)..(9)
;   OTHER INFORMATION: x=any residue
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (5)..(7)
;   OTHER INFORMATION: x=any residue
US-10-477-526-7

  Query Match           100.0%;  Score 33;  DB 5;  Length 10;
  Best Local Similarity  100.0%;  Pred. No. 3.5e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              ||||||||||
Db          1 FXXGXXXCXG 10
```

```
RESULT 8
US-10-477-526-8
; Sequence 8, Application US/10477526
; Publication No. US20050009021A1
; GENERAL INFORMATION:
;  APPLICANT: Wisniewski, Jan
;  APPLICANT:  Ramshaw, Heather
;  APPLICANT:  Petkovich, Martin P.
;  TITLE OF INVENTION: A Retinoic Acid Metabolizing Cytochrome P450
;  FILE REFERENCE: 11812-91
;  CURRENT APPLICATION NUMBER: US/10/477,526
;  CURRENT FILING DATE:  2003-11-21
;  PRIOR APPLICATION NUMBER: US 60/292,531
;  PRIOR FILING DATE: 2001-05-24
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-477-526-8

  Query Match             100.0%;  Score 33;  DB 5;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 3.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGGGARSCLG 10


RESULT 9
US-10-751-235-14
; Sequence 14, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
;  APPLICANT: DellaPenna, Dean
;  APPLICANT:  Tian, Li
;  APPLICANT:  Kim, Joonyul
;  TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
;  TITLE OF INVENTION:  Metabolism in Plants
;  FILE REFERENCE: MSU-08604
;  CURRENT APPLICATION NUMBER: US/10/751,235
;  CURRENT FILING DATE:  2004-01-02
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (2)..(3)
;   OTHER INFORMATION: Xaa can be any naturally occurring amino acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5)..(7)
;   OTHER INFORMATION: Xaa can be any naturally occurring amino acid
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (9)..(9)
;   OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-751-235-14

  Query Match             100.0%;  Score 33;  DB 5;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              ||||||||||
Db          1 FXXGXXXCXG 10


RESULT 10
US-10-751-235-15
; Sequence 15, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
;  APPLICANT: DellaPenna, Dean
;  APPLICANT:  Tian, Li
;  APPLICANT:  Kim, Joonyul
;  TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
;  TITLE OF INVENTION:  Metabolism in Plants
;  FILE REFERENCE: MSU-08604
;  CURRENT APPLICATION NUMBER: US/10/751,235
;  CURRENT FILING DATE:  2004-01-02
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 15
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-10-751-235-15

  Query Match             100.0%;  Score 33;  DB 5;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 3.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FSGGPRKCVG 10


RESULT 11
US-10-097-559-28
; Sequence 28, Application US/10097559
; Publication No. US20030166255A1
; GENERAL INFORMATION:
;  APPLICANT: Chappell, Joseph
;  APPLICANT:  Ralston, Lyle F.
;  TITLE OF INVENTION: Cytochrome P450s and Uses Thereof
;  FILE REFERENCE: 07678/100003
;  CURRENT APPLICATION NUMBER: US/10/097,559
;  CURRENT FILING DATE:  2002-03-08
;  PRIOR APPLICATION NUMBER: US 60/274,241
;  PRIOR FILING DATE: 2001-03-09
;  PRIOR APPLICATION NUMBER: US 60/275,597
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 28
;   LENGTH: 11
;   TYPE: PRT
;   ORGANISM: Nicotiana tabacum p450 protein
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 4,8
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-10-097-559-28

  Query Match             100.0%;  Score 33;  DB 4;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 3.8e+02;
  Matches    6;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |:||::||:|
Db          2 FGXGRRXCPG 11


RESULT 12
US-10-804-772-26
; Sequence 26, Application US/10804772
```

```
; Publication No. US20040244077A1
; GENERAL INFORMATION:
;  APPLICANT: Azpiroz, Ricardo
;  APPLICANT:  Choe, Sunghwa
;  APPLICANT:  Feldmann, Kenneth A.
;  TITLE OF INVENTION: DWF4 POLYNUCLEOTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: 11696-070001
;  CURRENT APPLICATION NUMBER: US/10/804,772
;  CURRENT FILING DATE:  2004-03-18
;  PRIOR APPLICATION NUMBER: US/09/502,426
;  PRIOR FILING DATE: 2000-02-11
;  PRIOR APPLICATION NUMBER: US 60/119,657
;  PRIOR FILING DATE: 1999-02-11
;  PRIOR APPLICATION NUMBER: US 60/119,658
;  PRIOR FILING DATE: 1999-02-11
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 26
;   LENGTH: 11
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Heme binding domain
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 4
;   OTHER INFORMATION: Xaa = Ala, Ser, or Val
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 8
;   OTHER INFORMATION: Xaa = Any Amino Acid
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 10
;   OTHER INFORMATION: Xaa = Pro, Ala, or Val
US-10-804-772-26

  Query Match             100.0%;  Score 33;  DB 5;  Length 11;
  Best Local Similarity   70.0%;  Pred. No. 3.8e+02;
  Matches    7;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |:||::||||
Db          2 FGXGRRXCXG 11


RESULT 13
US-10-804-772-29
; Sequence 29, Application US/10804772
; Publication No. US20040244077A1
; GENERAL INFORMATION:
;  APPLICANT: Azpiroz, Ricardo
;  APPLICANT:  Choe, Sunghwa
;  APPLICANT:  Feldmann, Kenneth A.
;  TITLE OF INVENTION: DWF4 POLYNUCLEOTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: 11696-070001
;  CURRENT APPLICATION NUMBER: US/10/804,772
;  CURRENT FILING DATE:  2004-03-18
;  PRIOR APPLICATION NUMBER: US/09/502,426
;  PRIOR FILING DATE: 2000-02-11
;  PRIOR APPLICATION NUMBER: US 60/119,657
;  PRIOR FILING DATE: 1999-02-11
;  PRIOR APPLICATION NUMBER: US 60/119,658
;  PRIOR FILING DATE: 1999-02-11
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29
;   LENGTH: 11
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Exemplary sequence
US-10-804-772-29

  Query Match             100.0%;  Score 33;  DB 5;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 3.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGGGPRLCAG 11


RESULT 14
US-10-373-877-10
; Sequence 10, Application US/10373877
; Publication No. US20030198986A1
;   GENERAL INFORMATION:
;        APPLICANT: Lathe, Richard
;                   Rose, Kenneth A.
;                   Stapleton, Genevieve
;        TITLE OF INVENTION: HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;                            SEQUENCES CODING THEREFOR AND USED THEREOF
;        NUMBER OF SEQUENCES: 45
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: NIXON & VANDERHYE P.C.
;             STREET: 1100 No. US20030198986A1th Glebe Rd.  8th floor
;             CITY: Arlington
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22201-4741
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/373,877
;             FILING DATE: 27-Feb-2003
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: PCT/GB95/02465
;             FILING DATE: 18-OCT-1995
;             APPLICATION NUMBER: 09/270,751
;             FILING DATE: 17-MAR-1999
;             APPLICATION NUMBER: 08/845,161
;             FILING DATE: 21-APR-1997
;             APPLICATION NUMBER: GB 9421093.7
;             FILING DATE: 19-OCT-1994
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Mitchard, Leonard C.
;             REGISTRATION NUMBER: 29,009
;             REFERENCE/DOCKET NUMBER: 604-572
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-816-4000
;             TELEFAX: 703-816-4100
;   INFORMATION FOR SEQ ID NO: 10:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 14 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: <Unknown>
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-373-877-10

  Query Match             100.0%;  Score 33;  DB 4;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 4.7e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              ||||||||||
Db          1 FXXGXXXCXG 10


RESULT 15
```

```
US-10-373-877-19
; Sequence 19, Application US/10373877
; Publication No. US20030198986A1
;   GENERAL INFORMATION:
;        APPLICANT: Lathe, Richard
;                   Rose, Kenneth A.
;                   Stapleton, Genevieve
;        TITLE OF INVENTION: HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;                            SEQUENCES CODING THEREFOR AND USED THEREOF
;        NUMBER OF SEQUENCES: 45
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: NIXON & VANDERHYE P.C.
;             STREET: 1100 No. US20030198986A1th Glebe Rd.  8th floor
;             CITY: Arlington
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22201-4741
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/373,877
;             FILING DATE: 27-Feb-2003
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: PCT/GB95/02465
;             FILING DATE: 18-OCT-1995
;             APPLICATION NUMBER: 09/270,751
;             FILING DATE: 17-MAR-1999
;             APPLICATION NUMBER: 08/845,161
;             FILING DATE: 21-APR-1997
;             APPLICATION NUMBER: GB 9421093.7
;             FILING DATE: 19-OCT-1994
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Mitchard, Leonard C.
;             REGISTRATION NUMBER: 29,009
;             REFERENCE/DOCKET NUMBER: 604-572
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-816-4000
;             TELEFAX: 703-816-4100
;   INFORMATION FOR SEQ ID NO: 19:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 14 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: <Unknown>
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-10-373-877-19

  Query Match             100.0%;  Score 33;  DB 4;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 4.7e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGLGTSKCPG 10


RESULT 16
US-10-373-877-20
; Sequence 20, Application US/10373877
; Publication No. US20030198986A1
;   GENERAL INFORMATION:
;        APPLICANT: Lathe, Richard
;                   Rose, Kenneth A.
;                   Stapleton, Genevieve
;        TITLE OF INVENTION: HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;                            SEQUENCES CODING THEREFOR AND USED THEREOF
;        NUMBER OF SEQUENCES: 45
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: NIXON & VANDERHYE P.C.
;             STREET: 1100 No. US20030198986A1th Glebe Rd.  8th floor
;             CITY: Arlington
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22201-4741
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/373,877
;             FILING DATE: 27-Feb-2003
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: PCT/GB95/02465
;             FILING DATE: 18-OCT-1995
;             APPLICATION NUMBER: 09/270,751
;             FILING DATE: 17-MAR-1999
;             APPLICATION NUMBER: 08/845,161
;             FILING DATE: 21-APR-1997
;             APPLICATION NUMBER: GB 9421093.7
;             FILING DATE: 19-OCT-1994
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Mitchard, Leonard C.
;             REGISTRATION NUMBER: 29,009
;             REFERENCE/DOCKET NUMBER: 604-572
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-816-4000
;             TELEFAX: 703-816-4100
;   INFORMATION FOR SEQ ID NO: 20:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 14 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: <Unknown>
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-10-373-877-20

  Query Match             100.0%;  Score 33;  DB 4;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 4.7e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGSGATICPG 10


RESULT 17
US-10-373-877-21
; Sequence 21, Application US/10373877
; Publication No. US20030198986A1
;   GENERAL INFORMATION:
;        APPLICANT: Lathe, Richard
;                   Rose, Kenneth A.
;                   Stapleton, Genevieve
;        TITLE OF INVENTION: HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;                            SEQUENCES CODING THEREFOR AND USED THEREOF
;        NUMBER OF SEQUENCES: 45
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: NIXON & VANDERHYE P.C.
;             STREET: 1100 No. US20030198986A1th Glebe Rd.  8th floor
;             CITY: Arlington
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22201-4741
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/10/373,877
;              FILING DATE: 27-Feb-2003
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: PCT/GB95/02465
;              FILING DATE: 18-OCT-1995
;              APPLICATION NUMBER: 09/270,751
;              FILING DATE: 17-MAR-1999
;              APPLICATION NUMBER: 08/845,161
;              FILING DATE: 21-APR-1997
;              APPLICATION NUMBER: GB 9421093.7
;              FILING DATE: 19-OCT-1994
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Mitchard, Leonard C.
;              REGISTRATION NUMBER: 29,009
;              REFERENCE/DOCKET NUMBER: 604-572
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 703-816-4000
;              TELEFAX: 703-816-4100
;    INFORMATION FOR SEQ ID NO: 21:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 14 amino acids
;              TYPE: amino acid
;              STRANDEDNESS: <Unknown>
;              TOPOLOGY: linear
;         MOLECULE TYPE: peptide
;         SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-373-877-21

  Query Match             100.0%;  Score 33;  DB 4;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 4.7e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGAGPRSCVG 10


RESULT 18
US-10-373-877-22
; Sequence 22, Application US/10373877
; Publication No. US20030198986A1
;   GENERAL INFORMATION:
;         APPLICANT: Lathe, Richard
;                    Rose, Kenneth A.
;                    Stapleton, Genevieve
;         TITLE OF INVENTION: HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;                             SEQUENCES CODING THEREFOR AND USED THEREOF
;         NUMBER OF SEQUENCES: 45
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: NIXON & VANDERHYE P.C.
;              STREET: 1100 No. US20030198986A1th Glebe Rd.  8th floor
;              CITY: Arlington
;              STATE: VA
;              COUNTRY: USA
;              ZIP: 22201-4741
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.30
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/10/373,877
;              FILING DATE: 27-Feb-2003
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: PCT/GB95/02465
;              FILING DATE: 18-OCT-1995
;              APPLICATION NUMBER: 09/270,751
;              FILING DATE: 17-MAR-1999
;              APPLICATION NUMBER: 08/845,161
;              FILING DATE: 21-APR-1997
;              APPLICATION NUMBER: GB 9421093.7
;              FILING DATE: 19-OCT-1994
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Mitchard, Leonard C.
;              REGISTRATION NUMBER: 29,009
;              REFERENCE/DOCKET NUMBER: 604-572
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 703-816-4000
;              TELEFAX: 703-816-4100
;    INFORMATION FOR SEQ ID NO: 22:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 14 amino acids
;              TYPE: amino acid
;              STRANDEDNESS: <Unknown>
;              TOPOLOGY: linear
;         MOLECULE TYPE: peptide
;         SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-10-373-877-22

  Query Match             100.0%;  Score 33;  DB 4;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 4.7e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 PGFGMRQCLG 10


RESULT 19
US-10-373-877-23
; Sequence 23, Application US/10373877
; Publication No. US20030198986A1
;   GENERAL INFORMATION:
;         APPLICANT: Lathe, Richard
;                    Rose, Kenneth A.
;                    Stapleton, Genevieve
;         TITLE OF INVENTION: HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;                             SEQUENCES CODING THEREFOR AND USED THEREOF
;         NUMBER OF SEQUENCES: 45
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: NIXON & VANDERHYE P.C.
;              STREET: 1100 No. US20030198986A1th Glebe Rd.  8th floor
;              CITY: Arlington
;              STATE: VA
;              COUNTRY: USA
;              ZIP: 22201-4741
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.30
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/10/373,877
;              FILING DATE: 27-Feb-2003
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: PCT/GB95/02465
;              FILING DATE: 18-OCT-1995
;              APPLICATION NUMBER: 09/270,751
;              FILING DATE: 17-MAR-1999
;              APPLICATION NUMBER: 08/845,161
;              FILING DATE: 21-APR-1997
;              APPLICATION NUMBER: GB 9421093.7
;              FILING DATE: 19-OCT-1994
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Mitchard, Leonard C.
;              REGISTRATION NUMBER: 29,009
;              REFERENCE/DOCKET NUMBER: 604-572
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 703-816-4000
;              TELEFAX: 703-816-4100
;    INFORMATION FOR SEQ ID NO: 23:
;         SEQUENCE CHARACTERISTICS:
```

```
;             LENGTH: 14 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: <Unknown>
;             TOPOLOGY: linear
;          MOLECULE TYPE: peptide
;          SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-10-373-877-23

  Query Match             100.0%;  Score 33;  DB 4;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 4.7e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGCGARVCLG 10


RESULT 20
US-10-373-877-24
; Sequence 24, Application US/10373877
; Publication No. US20030198986A1
;   GENERAL INFORMATION:
;        APPLICANT: Lathe, Richard
;                   Rose, Kenneth A.
;                   Stapleton, Genevieve
;        TITLE OF INVENTION: HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;                            SEQUENCES CODING THEREFOR AND USED THEREOF
;        NUMBER OF SEQUENCES: 45
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: NIXON & VANDERHYE P.C.
;             STREET: 1100 No. US20030198986A1th Glebe Rd.  8th floor
;             CITY: Arlington
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22201-4741
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/373,877
;             FILING DATE: 27-Feb-2003
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: PCT/GB95/02465
;             FILING DATE: 18-OCT-1995
;             APPLICATION NUMBER: 09/270,751
;             FILING DATE: 17-MAR-1999
;             APPLICATION NUMBER: 08/845,161
;             FILING DATE: 21-APR-1997
;             APPLICATION NUMBER: GB 9421093.7
;             FILING DATE: 19-OCT-1994
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Mitchard, Leonard C.
;             REGISTRATION NUMBER: 29,009
;             REFERENCE/DOCKET NUMBER: 604-572
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-816-4000
;             TELEFAX: 703-816-4100
;   INFORMATION FOR SEQ ID NO: 24:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 14 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: <Unknown>
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-10-373-877-24

  Query Match             100.0%;  Score 33;  DB 4;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 4.7e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGWGVRQCLG 10


RESULT 21
US-10-373-877-25
; Sequence 25, Application US/10373877
; Publication No. US20030198986A1
;   GENERAL INFORMATION:
;        APPLICANT: Lathe, Richard
;                   Rose, Kenneth A.
;                   Stapleton, Genevieve
;        TITLE OF INVENTION: HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;                            SEQUENCES CODING THEREFOR AND USED THEREOF
;        NUMBER OF SEQUENCES: 45
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: NIXON & VANDERHYE P.C.
;             STREET: 1100 No. US20030198986A1th Glebe Rd.  8th floor
;             CITY: Arlington
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22201-4741
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/373,877
;             FILING DATE: 27-Feb-2003
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: PCT/GB95/02465
;             FILING DATE: 18-OCT-1995
;             APPLICATION NUMBER: 09/270,751
;             FILING DATE: 17-MAR-1999
;             APPLICATION NUMBER: 08/845,161
;             FILING DATE: 21-APR-1997
;             APPLICATION NUMBER: GB 9421093.7
;             FILING DATE: 19-OCT-1994
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Mitchard, Leonard C.
;             REGISTRATION NUMBER: 29,009
;             REFERENCE/DOCKET NUMBER: 604-572
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-816-4000
;             TELEFAX: 703-816-4100
;   INFORMATION FOR SEQ ID NO: 25:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 14 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: <Unknown>
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-10-373-877-25

  Query Match             100.0%;  Score 33;  DB 4;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 4.7e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGYGVRACLG 10


RESULT 22
US-10-865-478-774
; Sequence 774, Application US/10865478
; Publication No. US20040235041A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Shimkets, Richard A.
;  APPLICANT:  Leach, Martin D.
;  TITLE OF INVENTION: cSingle Nucleotide Polymorphisms for Known Genes
;  FILE REFERENCE: 15966-534-CIP1
;  CURRENT APPLICATION NUMBER: US/10/865,478
;  CURRENT FILING DATE:  2004-06-10
;  PRIOR APPLICATION NUMBER: 60/109,024
;  PRIOR FILING DATE: 1998-11-17
;  PRIOR APPLICATION NUMBER: 09/443,199
;  PRIOR FILING DATE: 1999-11-16
;  PRIOR APPLICATION NUMBER: 09/442,129
;  PRIOR FILING DATE: 1999-11-16
;  PRIOR APPLICATION NUMBER: 09/442,849
;  PRIOR FILING DATE: 1999-11-17
;  NUMBER OF SEQ ID NOS: 880
;  SOFTWARE: CuraGen Patent Formatter Version 0.9
; SEQ ID NO 774
;   LENGTH: 14
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (7)...(0)
;   OTHER INFORMATION: cSNP translation
US-10-865-478-774

  Query Match           100.0%;  Score 33;  DB 5;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 4.7e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FHYGVLACEG 11


RESULT 23
US-09-957-674-13
; Sequence 13, Application US/09957674
; Patent No. US20020120948A1
; GENERAL INFORMATION:
;  APPLICANT: Medical Research Council
;  TITLE OF INVENTION: Methods for Expressing Gene Products
;  FILE REFERENCE: 18396/2072
;  CURRENT APPLICATION NUMBER: US/09/957,674
;  CURRENT FILING DATE:  2001-09-20
;  PRIOR APPLICATION NUMBER: GB990736
;  PRIOR FILING DATE: 1999-03-30
;  PRIOR APPLICATION NUMBER: PCT/GB00/01225
;  PRIOR FILING DATE: 2000-03-30
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Hepatitis B virus
US-09-957-674-13

  Query Match           100.0%;  Score 33;  DB 3;  Length 15;
  Best Local Similarity   40.0%;  Pred. No. 5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          4 FLGGITVCLG 13


RESULT 24
US-10-257-378-1
; Sequence 1, Application US/10257378
; Publication No. US20030190642A1
; GENERAL INFORMATION:
;  APPLICANT: Jones, Glenville
;  APPLICANT:  Petkovich, P. Martin
;  APPLICANT:  White, Jay
;  APPLICANT:  Ramshaw, Heather A.
;  APPLICANT:  Stangle, Wayne A.
;  TITLE OF INVENTION: A Thymus Expressed Human Cytochrome P450 (P450TEC)
;  FILE REFERENCE: 11812-65
;  CURRENT APPLICATION NUMBER: US/10/257,378
;  CURRENT FILING DATE:  2001-04-20
;  PRIOR APPLICATION NUMBER: US 60/208,785
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: US 60/198,617
;  PRIOR FILING DATE: 2000-04-20
;  NUMBER OF SEQ ID NOS: 56
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: query sequence
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (3)
;   OTHER INFORMATION: Xaa can be Gly or Ser
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (4)
;   OTHER INFORMATION: Xaa can be any amino acid
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (6)
;   OTHER INFORMATION: Xaa can be any amino  acid
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (8)
;   OTHER INFORMATION: Xaa can be any amino acid
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (7)
;   OTHER INFORMATION: Xaa can be Ala, Arg, or His
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (10)
;   OTHER INFORMATION: Xaa can be any amino acid
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (12)..(13)
;   OTHER INFORMATION: Xaa can be amino  acid
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (14)
;   OTHER INFORMATION: Xaa can be Phe, Leu,  or Ile
US-10-257-378-1

  Query Match           100.0%;  Score 33;  DB 4;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 5e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              ||||||||||
Db          2 FXXGXXXCXG 11


RESULT 25
US-10-882-241-17
; Sequence 17, Application US/10882241
; Publication No. US20050114910A1
; GENERAL INFORMATION:
;  APPLICANT: LONE, YU-CHUN
;  APPLICANT:  PAJOT, ANTHONY
;  APPLICANT:  AURIAULT, CLAUDE
;  APPLICANT:  PANCRE, VERONIQUE
;  APPLICANT:  LEMONNIER, FRANCOIS
;  TITLE OF INVENTION: TRANSGENIC MICE HAVING A HUMAN MAJOR HISTOCOMPATIBILITY
```

```
;  TITLE OF INVENTION:  COMPLEX (MHC) PHENOTYPE, EXPERIMENTAL USES AND APPLICATIONS
;  FILE REFERENCE: 03495.0319
;  CURRENT APPLICATION NUMBER: US/10/882,241
;  CURRENT FILING DATE:  2004-07-02
;  PRIOR APPLICATION NUMBER: 60/490,945
;  PRIOR FILING DATE: 2003-08-05
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 17
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: Peptide
US-10-882-241-17

  Query Match             100.0%;  Score 33;  DB 5;  Length 15;
  Best Local Similarity   40.0%;  Pred. No. 5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          5 FLGGTTVCLG 14


RESULT 26
US-10-163-198-90
; Sequence 90, Application US/10163198
; Publication No. US20030126645A1
; GENERAL INFORMATION:
;  APPLICANT: Rebecca E. Cahoon
;  APPLICANT:  Elmer P. Heppard
;  APPLICANT:  No. US20030126645A1uhiro Nagasawa
;  APPLICANT:  Hajime Sakai
;  TITLE OF INVENTION: Alteration Of Embryo/Endosperm Size During Seed Development
;  FILE REFERENCE: BB1487 US NA
;  CURRENT APPLICATION NUMBER: US/10/163,198
;  CURRENT FILING DATE:  2002-06-05
;  PRIOR APPLICATION NUMBER: 60/295,921
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: 60/334,317
;  PRIOR FILING DATE: 2001-11-28
;  NUMBER OF SEQ ID NOS: 111
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 90
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: conserved sequence motif
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (9)..(9)
;   OTHER INFORMATION: Xaa = any amino acid
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (13)..(13)
;   OTHER INFORMATION: Xaa = any amino acid
US-10-163-198-90

  Query Match             100.0%;  Score 33;  DB 4;  Length 16;
  Best Local Similarity   60.0%;  Pred. No. 5.3e+02;
  Matches    6;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |:||::||:|
Db          7 FGXGRRXCPG 16


RESULT 27
US-10-926-683-1257
; Sequence 1257, Application US/10926683
; Publication No. US20050106595A1
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.025CP1
;  CURRENT APPLICATION NUMBER: US/10/926,683
;  CURRENT FILING DATE:  2004-08-25
;  PRIOR APPLICATION NUMBER: US/09/471,276
;  PRIOR FILING DATE: 1999-12-21
;  PRIOR APPLICATION NUMBER: 09/057,719
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: 09/069,047
;  PRIOR FILING DATE: 1998-04-28
;  PRIOR APPLICATION NUMBER: PCT/IB99/00712
;  PRIOR FILING DATE: 1999-04-09
;  NUMBER OF SEQ ID NOS: 1622
;  SOFTWARE: Patent.pm
; SEQ ID NO 1257
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIGNAL
;   LOCATION: -14..-1
US-10-926-683-1257

  Query Match             100.0%;  Score 33;  DB 5;  Length 16;
  Best Local Similarity   40.0%;  Pred. No. 5.3e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          4 FFFGNSPCCG 13


RESULT 28
US-08-647-444-12
; Sequence 12, Application US/08647444
; Publication No. US20020127548A1
;  GENERAL INFORMATION:
;     APPLICANT:  Seidman, Christine
;     APPLICANT:  Seidman, Jonathan
;     APPLICANT:  Thierfelder, Ludwig
;     APPLICANT:  Watkins, Hugh
;     APPLICANT:  Mcrae, Calum
;     TITLE OF INVENTION:  Methods for Detecting Mutations Associated
;     TITLE OF INVENTION:  with Hypertrophic Cardiomyopathy
;     NUMBER OF SEQUENCES:  12
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  LAHIVE & COCKFIELD
;       STREET:  60 State Street, suite 510
;       CITY:  Boston
;       STATE:  Massachusetts
;       COUNTRY:  USA
;       ZIP:  02109-1875
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/647,444
;       FILING DATE:  11-NOV-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Elizabeth A. Hanley
;       REGISTRATION NUMBER:  33,505
;       REFERENCE/DOCKET NUMBER:  IGI-037CP
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617)227-7400
```

```
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  18 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-647-444-12

  Query Match           100.0%;  Score 33;  DB 2;  Length 18;
  Best Local Similarity  40.0%;  Pred. No. 5.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FDGGIYVCGG 11


RESULT 29
US-10-855-595-6
; Sequence 6, Application US/10855595
; Publication No. US20040235057A1
;   GENERAL INFORMATION:
;        APPLICANT: Petkovich, P. Martin, White, Jay A.,
;                   Beckett, Barbara R., Jones, Glenville
;        TITLE OF INVENTION: Retinoid Metabolizing Protein
;        NUMBER OF SEQUENCES: 43
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Blake, Cassels & Graydon
;             STREET: Box 25, Commerce Court West
;             CITY: Toronto
;             STATE: Ontario
;             COUNTRY: Canada
;             ZIP: M5L 1A9
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
;             COMPUTER: COMPAQ, IBM PC compatible
;             OPERATING SYSTEM: MS-DOS 5.1
;             SOFTWARE: WORD PERFECT
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/855,595
;             FILING DATE: 28-May-2004
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/668,482
;             FILING DATE: 25-Sep-2000
;             APPLICATION NUMBER: 08/882,164
;             FILING DATE: June 25, 1997
;             APPLICATION NUMBER: 08/667,546
;             FILING DATE: June 21, 1996
;             APPLICATION NUMBER: 08/724,466
;             FILING DATE: October 1, 1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Hunt, John C.
;             REGISTRATION NUMBER: 36,424
;             REFERENCE/DOCKET NUMBER: 50767/00010
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (416) 863-4344
;             TELEFAX: (416) 863-2653
;   INFORMATION FOR SEQ ID NO: 6
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 20 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 6
US-10-855-595-6

  Query Match           100.0%;  Score 33;  DB 5;  Length 20;
  Best Local Similarity  40.0%;  Pred. No. 6.6e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGGGPRLCPG 11


RESULT 30
US-10-855-595-7
; Sequence 7, Application US/10855595
; Publication No. US20040235057A1
;   GENERAL INFORMATION:
;        APPLICANT: Petkovich, P. Martin, White, Jay A.,
;                   Beckett, Barbara R., Jones, Glenville
;        TITLE OF INVENTION: Retinoid Metabolizing Protein
;        NUMBER OF SEQUENCES: 43
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Blake, Cassels & Graydon
;             STREET: Box 25, Commerce Court West
;             CITY: Toronto
;             STATE: Ontario
;             COUNTRY: Canada
;             ZIP: M5L 1A9
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
;             COMPUTER: COMPAQ, IBM PC compatible
;             OPERATING SYSTEM: MS-DOS 5.1
;             SOFTWARE: WORD PERFECT
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/855,595
;             FILING DATE: 28-May-2004
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/668,482
;             FILING DATE: 25-Sep-2000
;             APPLICATION NUMBER: 08/882,164
;             FILING DATE: June 25, 1997
;             APPLICATION NUMBER: 08/667,546
;             FILING DATE: June 21, 1996
;             APPLICATION NUMBER: 08/724,466
;             FILING DATE: October 1, 1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Hunt, John C.
;             REGISTRATION NUMBER: 36,424
;             REFERENCE/DOCKET NUMBER: 50767/00010
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (416) 863-4344
;             TELEFAX: (416) 863-2653
;   INFORMATION FOR SEQ ID NO: 7
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 20 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 7
US-10-855-595-7

  Query Match           100.0%;  Score 33;  DB 5;  Length 20;
  Best Local Similarity  40.0%;  Pred. No. 6.6e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FSGGARNCIG 11


RESULT 31
US-10-855-595-8
; Sequence 8, Application US/10855595
; Publication No. US20040235057A1
;   GENERAL INFORMATION:
;        APPLICANT: Petkovich, P. Martin, White, Jay A.,
;                   Beckett, Barbara R., Jones, Glenville
;        TITLE OF INVENTION: Retinoid Metabolizing Protein
;        NUMBER OF SEQUENCES: 43
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Blake, Cassels & Graydon
;             STREET: Box 25, Commerce Court West
```

```
;             CITY: Toronto
;             STATE: Ontario
;             COUNTRY: Canada
;             ZIP: M5L 1A9
;       COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
;             COMPUTER: COMPAQ, IBM PC compatible
;             OPERATING SYSTEM: MS-DOS 5.1
;             SOFTWARE: WORD PERFECT
;       CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/855,595
;             FILING DATE: 28-May-2004
;       PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/668,482
;             FILING DATE: 25-Sep-2000
;             APPLICATION NUMBER: 08/882,164
;             FILING DATE: June 25, 1997
;             APPLICATION NUMBER: 08/667,546
;             FILING DATE: June 21, 1996
;             APPLICATION NUMBER: 08/724,466
;             FILING DATE: October 1, 1996
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Hunt, John C.
;             REGISTRATION NUMBER: 36,424
;             REFERENCE/DOCKET NUMBER: 50767/00010
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (416) 863-4344
;             TELEFAX: (416) 863-2653
;   INFORMATION FOR SEQ ID NO: 8
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 20 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 8
US-10-855-595-8

  Query Match             100.0%;  Score 33;  DB 5;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 6.6e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FSGGARNCIG 11


RESULT 32
US-10-855-595-9
; Sequence 9, Application US/10855595
; Publication No. US20040235057A1
;   GENERAL INFORMATION:
;       APPLICANT: Petkovich, P. Martin, White, Jay A.,
;                  Beckett, Barbara R., Jones, Glenville
;       TITLE OF INVENTION: Retinoid Metabolizing Protein
;       NUMBER OF SEQUENCES: 43
;       CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Blake, Cassels & Graydon
;             STREET: Box 25, Commerce Court West
;             CITY: Toronto
;             STATE: Ontario
;             COUNTRY: Canada
;             ZIP: M5L 1A9
;       COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
;             COMPUTER: COMPAQ, IBM PC compatible
;             OPERATING SYSTEM: MS-DOS 5.1
;             SOFTWARE: WORD PERFECT
;       CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/855,595
;             FILING DATE: 28-May-2004
;       PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/668,482
;             FILING DATE: 25-Sep-2000
;             APPLICATION NUMBER: 08/882,164
;             FILING DATE: June 25, 1997
;             APPLICATION NUMBER: 08/667,546
;             FILING DATE: June 21, 1996
;             APPLICATION NUMBER: 08/724,466
;             FILING DATE: October 1, 1996
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Hunt, John C.
;             REGISTRATION NUMBER: 36,424
;             REFERENCE/DOCKET NUMBER: 50767/00010
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (416) 863-4344
;             TELEFAX: (416) 863-2653
;   INFORMATION FOR SEQ ID NO: 9
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 20 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 9
US-10-855-595-9

  Query Match             100.0%;  Score 33;  DB 5;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 6.6e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGTGPRNCIG 11


RESULT 33
US-10-855-595-10
; Sequence 10, Application US/10855595
; Publication No. US20040235057A1
;   GENERAL INFORMATION:
;       APPLICANT: Petkovich, P. Martin, White, Jay A.,
;                  Beckett, Barbara R., Jones, Glenville
;       TITLE OF INVENTION: Retinoid Metabolizing Protein
;       NUMBER OF SEQUENCES: 43
;       CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Blake, Cassels & Graydon
;             STREET: Box 25, Commerce Court West
;             CITY: Toronto
;             STATE: Ontario
;             COUNTRY: Canada
;             ZIP: M5L 1A9
;       COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
;             COMPUTER: COMPAQ, IBM PC compatible
;             OPERATING SYSTEM: MS-DOS 5.1
;             SOFTWARE: WORD PERFECT
;       CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/855,595
;             FILING DATE: 28-May-2004
;       PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/668,482
;             FILING DATE: 25-Sep-2000
;             APPLICATION NUMBER: 08/882,164
;             FILING DATE: June 25, 1997
;             APPLICATION NUMBER: 08/667,546
;             FILING DATE: June 21, 1996
;             APPLICATION NUMBER: 08/724,466
;             FILING DATE: October 1, 1996
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Hunt, John C.
;             REGISTRATION NUMBER: 36,424
;             REFERENCE/DOCKET NUMBER: 50767/00010
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (416) 863-4344
;             TELEFAX: (416) 863-2653
;   INFORMATION FOR SEQ ID NO: 10
;       SEQUENCE CHARACTERISTICS:
```

```
;            LENGTH: 20 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 10
US-10-855-595-10

  Query Match             100.0%;  Score 33;  DB 5;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 6.6e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FSGGSRNCIG 11


RESULT 34
US-10-855-532-6
; Sequence 6, Application US/10855532
; Publication No. US20040259074A1
;   GENERAL INFORMATION:
;       APPLICANT: Petkovich, P. Martin, White, Jay A.,
;                  Beckett, Barbara R., Jones, Glenville
;       TITLE OF INVENTION: Retinoid Metabolizing Protein
;       NUMBER OF SEQUENCES: 43
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Blake, Cassels & Graydon
;            STREET: Box 25, Commerce Court West
;            CITY: Toronto
;            STATE: Ontario
;            COUNTRY: Canada
;            ZIP: M5L 1A9
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
;            COMPUTER: COMPAQ, IBM PC compatible
;            OPERATING SYSTEM: MS-DOS 5.1
;            SOFTWARE: WORD PERFECT
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/855,532
;            FILING DATE: 28-May-2004
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/668,482
;            FILING DATE: 25-Sep-2000
;            APPLICATION NUMBER: 08/882,164
;            FILING DATE: June 25, 1997
;            APPLICATION NUMBER: 08/667,546
;            FILING DATE: June 21, 1996
;            APPLICATION NUMBER: 08/724,466
;            FILING DATE: October 1, 1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Hunt, John C.
;            REGISTRATION NUMBER: 36,424
;            REFERENCE/DOCKET NUMBER: 50767/00010
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (416) 863-4344
;            TELEFAX: (416) 863-2653
;   INFORMATION FOR SEQ ID NO: 6
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 20 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 6
US-10-855-532-6

  Query Match             100.0%;  Score 33;  DB 5;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 6.6e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGGGPRLCPG 11


RESULT 35
US-10-855-532-7
; Sequence 7, Application US/10855532
; Publication No. US20040259074A1
;   GENERAL INFORMATION:
;       APPLICANT: Petkovich, P. Martin, White, Jay A.,
;                  Beckett, Barbara R., Jones, Glenville
;       TITLE OF INVENTION: Retinoid Metabolizing Protein
;       NUMBER OF SEQUENCES: 43
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Blake, Cassels & Graydon
;            STREET: Box 25, Commerce Court West
;            CITY: Toronto
;            STATE: Ontario
;            COUNTRY: Canada
;            ZIP: M5L 1A9
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
;            COMPUTER: COMPAQ, IBM PC compatible
;            OPERATING SYSTEM: MS-DOS 5.1
;            SOFTWARE: WORD PERFECT
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/855,532
;            FILING DATE: 28-May-2004
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/668,482
;            FILING DATE: 25-Sep-2000
;            APPLICATION NUMBER: 08/882,164
;            FILING DATE: June 25, 1997
;            APPLICATION NUMBER: 08/667,546
;            FILING DATE: June 21, 1996
;            APPLICATION NUMBER: 08/724,466
;            FILING DATE: October 1, 1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Hunt, John C.
;            REGISTRATION NUMBER: 36,424
;            REFERENCE/DOCKET NUMBER: 50767/00010
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (416) 863-4344
;            TELEFAX: (416) 863-2653
;   INFORMATION FOR SEQ ID NO: 7
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 20 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 7
US-10-855-532-7

  Query Match             100.0%;  Score 33;  DB 5;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 6.6e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FSGGARNCIG 11


RESULT 36
US-10-855-532-8
; Sequence 8, Application US/10855532
; Publication No. US20040259074A1
;   GENERAL INFORMATION:
;       APPLICANT: Petkovich, P. Martin, White, Jay A.,
;                  Beckett, Barbara R., Jones, Glenville
;       TITLE OF INVENTION: Retinoid Metabolizing Protein
;       NUMBER OF SEQUENCES: 43
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Blake, Cassels & Graydon
;            STREET: Box 25, Commerce Court West
;            CITY: Toronto
;            STATE: Ontario
```

```
;            COUNTRY: Canada
;            ZIP: M5L 1A9
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
;            COMPUTER: COMPAQ, IBM PC compatible
;            OPERATING SYSTEM: MS-DOS 5.1
;            SOFTWARE: WORD PERFECT
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/855,532
;            FILING DATE: 28-May-2004
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/668,482
;            FILING DATE: 25-Sep-2000
;            APPLICATION NUMBER: 08/882,164
;            FILING DATE: June 25, 1997
;            APPLICATION NUMBER: 08/667,546
;            FILING DATE: June 21, 1996
;            APPLICATION NUMBER: 08/724,466
;            FILING DATE: October 1, 1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Hunt, John C.
;            REGISTRATION NUMBER: 36,424
;            REFERENCE/DOCKET NUMBER: 50767/00010
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (416) 863-4344
;            TELEFAX: (416) 863-2653
;  INFORMATION FOR SEQ ID NO: 8
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 20 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 8
US-10-855-532-8

  Query Match             100.0%;  Score 33;  DB 5;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 6.6e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FSGGARNCIG 11


RESULT 37
US-10-855-532-9
; Sequence 9, Application US/10855532
; Publication No. US20040259074A1
;  GENERAL INFORMATION:
;       APPLICANT: Petkovich, P. Martin, White, Jay A.,
;                  Beckett, Barbara R., Jones, Glenville
;       TITLE OF INVENTION: Retinoid Metabolizing Protein
;       NUMBER OF SEQUENCES: 43
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Blake, Cassels & Graydon
;            STREET: Box 25, Commerce Court West
;            CITY: Toronto
;            STATE: Ontario
;            COUNTRY: Canada
;            ZIP: M5L 1A9
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
;            COMPUTER: COMPAQ, IBM PC compatible
;            OPERATING SYSTEM: MS-DOS 5.1
;            SOFTWARE: WORD PERFECT
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/855,532
;            FILING DATE: 28-May-2004
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/668,482
;            FILING DATE: 25-Sep-2000
;            APPLICATION NUMBER: 08/882,164
;            FILING DATE: June 25, 1997
;            APPLICATION NUMBER: 08/667,546
;            FILING DATE: June 21, 1996
;            APPLICATION NUMBER: 08/724,466
;            FILING DATE: October 1, 1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Hunt, John C.
;            REGISTRATION NUMBER: 36,424
;            REFERENCE/DOCKET NUMBER: 50767/00010
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (416) 863-4344
;            TELEFAX: (416) 863-2653
;  INFORMATION FOR SEQ ID NO: 9
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 20 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 9
US-10-855-532-9

  Query Match             100.0%;  Score 33;  DB 5;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 6.6e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGTGPRNCIG 11


RESULT 38
US-10-855-532-10
; Sequence 10, Application US/10855532
; Publication No. US20040259074A1
;  GENERAL INFORMATION:
;       APPLICANT: Petkovich, P. Martin, White, Jay A.,
;                  Beckett, Barbara R., Jones, Glenville
;       TITLE OF INVENTION: Retinoid Metabolizing Protein
;       NUMBER OF SEQUENCES: 43
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Blake, Cassels & Graydon
;            STREET: Box 25, Commerce Court West
;            CITY: Toronto
;            STATE: Ontario
;            COUNTRY: Canada
;            ZIP: M5L 1A9
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
;            COMPUTER: COMPAQ, IBM PC compatible
;            OPERATING SYSTEM: MS-DOS 5.1
;            SOFTWARE: WORD PERFECT
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/855,532
;            FILING DATE: 28-May-2004
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/668,482
;            FILING DATE: 25-Sep-2000
;            APPLICATION NUMBER: 08/882,164
;            FILING DATE: June 25, 1997
;            APPLICATION NUMBER: 08/667,546
;            FILING DATE: June 21, 1996
;            APPLICATION NUMBER: 08/724,466
;            FILING DATE: October 1, 1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Hunt, John C.
;            REGISTRATION NUMBER: 36,424
;            REFERENCE/DOCKET NUMBER: 50767/00010
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (416) 863-4344
;            TELEFAX: (416) 863-2653
;  INFORMATION FOR SEQ ID NO: 10
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 20 amino acids
;            TYPE: amino acid
```

```
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 10
US-10-855-532-10

  Query Match            100.0%;  Score 33;  DB 5;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 6.6e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FSGGSRNCIG 11


RESULT 39
US-10-021-425-25
; Sequence 25, Application US/10021425
; Publication No. US20030148420A1
; GENERAL INFORMATION:
;  APPLICANT: Suzanne L. Bolten
;  APPLICANT:  Alan M. Easton
;  APPLICANT:  Leslie C. Engel
;  APPLICANT:  Dean M. Messing
;  APPLICANT:  John S. Ng
;  APPLICANT:  Beverly A. Reitz
;  APPLICANT:  Scott A. Vaccaro
;  APPLICANT:  Mark C. Walker
;  APPLICANT:  Ping T. Wang
;  APPLICANT:  Robin A. Weinberg
;  TITLE OF INVENTION: Aspergillus ochraceus 11 alpha
;  TITLE OF INVENTION:  hydroxylase and oxidoreductase
;  FILE REFERENCE: S03196-00-US
;  CURRENT APPLICATION NUMBER: US/10/021,425
;  CURRENT FILING DATE:  2001-10-30
;  PRIOR APPLICATION NUMBER: USSN 60/244,300
;  PRIOR FILING DATE: 2000-10-30
;  NUMBER OF SEQ ID NOS: 65
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 25
;   LENGTH: 21
;   TYPE: PRT
;   ORGANISM: Aspergillus 11aOH peptide 3
US-10-021-425-25

  Query Match            100.0%;  Score 33;  DB 4;  Length 21;
  Best Local Similarity   40.0%;  Pred. No. 6.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          4 FGHGVHACPG 13


RESULT 40
US-10-900-856-28
; Sequence 28, Application US/10900856
; Publication No. US20050003473A1
; GENERAL INFORMATION:
;  APPLICANT: Bolten, Suzanne L
;  APPLICANT:  Leslie, Engel C
;  APPLICANT:  Dean, Messing M
;  APPLICANT:  John, Ng S
;  APPLICANT:  Beverly, Reitz A
;  APPLICANT:  Scott, Vaccaro A
;  APPLICANT:  Mark, Walker C
;  APPLICANT:  Ping, Wang T
;  APPLICANT:  Robin, Weinberg A
;  TITLE OF INVENTION: Aspergillus ochraceus 11 alpha hydroxylase and oxidoreductase
;  FILE REFERENCE: 3196
;  CURRENT APPLICATION NUMBER: US/10/900,856
;  CURRENT FILING DATE:  2004-07-28
;  NUMBER OF SEQ ID NOS: 68
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
;   LENGTH: 21
;   TYPE: PRT
;   ORGANISM: aspergillus ochraceus 11aOH peptide 3
US-10-900-856-28

  Query Match            100.0%;  Score 33;  DB 5;  Length 21;
  Best Local Similarity   40.0%;  Pred. No. 6.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          4 FGHGVHACPG 13


RESULT 41
US-10-103-196-31
; Sequence 31, Application US/10103196
; Publication No. US20030050466A1
; GENERAL INFORMATION:
;  APPLICANT: Ni et al.
;  TITLE OF INVENTION: TM4SF Polynucleotides, Polypeptides, and Antibodies
;  FILE REFERENCE: PT008P1
;  CURRENT APPLICATION NUMBER: US/10/103,196
;  CURRENT FILING DATE:  2002-03-22
;  PRIOR APPLICATION NUMBER: 09/707,936
;  PRIOR FILING DATE: 2000-11-08
;  PRIOR APPLICATION NUMBER: PCT/US00/13504
;  PRIOR FILING DATE: 2000-05-18
;  PRIOR APPLICATION NUMBER: 60/178,770
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: 60/149,447
;  PRIOR FILING DATE: 1999-08-18
;  PRIOR APPLICATION NUMBER: 60/138,573
;  PRIOR FILING DATE: 1999-06-11
;  PRIOR APPLICATION NUMBER: 60/137,797
;  PRIOR FILING DATE: 1999-06-03
;  PRIOR APPLICATION NUMBER: 60/135,122
;  PRIOR FILING DATE: 1999-05-19
;  NUMBER OF SEQ ID NOS: 44
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
;   LENGTH: 23
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-103-196-31

  Query Match            100.0%;  Score 33;  DB 4;  Length 23;
  Best Local Similarity   40.0%;  Pred. No. 7.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          5 FLLGFLGCCG 14


RESULT 42
US-09-749-637A-126
; Sequence 126, Application US/09749637A
; Patent No. US20020173449A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M.
;  APPLICANT:  Cartier, G. Edward
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Hillyard, David R.
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Layer, Richard T.
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: O-Superfamily Conotoxin Peptides
;  FILE REFERENCE: 2314-227
;  CURRENT APPLICATION NUMBER: US/09/749,637A
```

```
;  CURRENT FILING DATE:  2000-12-28
;  PRIOR APPLICATION NUMBER: US 60/243,412
;  PRIOR FILING DATE: 2000-10-27
;  PRIOR APPLICATION NUMBER: US60/219,440
;  PRIOR FILING DATE: 2000-07-20
;  PRIOR APPLICATION NUMBER: US 60/214,263
;  PRIOR FILING DATE: 2000-06-26
;  PRIOR APPLICATION NUMBER: US 60/173,754
;  PRIOR FILING DATE: 1999-12-30
;  NUMBER OF SEQ ID NOS: 409
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 126
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Conus marmoreus
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (1)..(30)
;   OTHER INFORMATION: Xaa at residues 3 and 5 may be Glu or gamma-carboxy-Glu; Xaa at r
;   OTHER INFORMATION: esidues 6, 7, 18 and 19 may be Pro or hydroxy-Pro; Xaa at residue
;   OTHER INFORMATION: s 24 and 28 may be Trp or bromo-Trp
US-09-749-637A-126

  Query Match             100.0%;  Score 33;  DB 3;  Length 30;
  Best Local Similarity   60.0%;  Pred. No. 9.5e+02;
  Matches    6;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|||:|:|
Db         14 FKIGXXCCSG 23


RESULT 43
US-09-749-637A-129
; Sequence 129, Application US/09749637A
; Patent No. US20020173449A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M.
;  APPLICANT:  Cartier, G. Edward
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Hillyard, David R.
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Layer, Richard T.
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: O-Superfamily Conotoxin Peptides
;  FILE REFERENCE: 2314-227
;  CURRENT APPLICATION NUMBER: US/09/749,637A
;  CURRENT FILING DATE:  2000-12-28
;  PRIOR APPLICATION NUMBER: US 60/243,412
;  PRIOR FILING DATE: 2000-10-27
;  PRIOR APPLICATION NUMBER: US60/219,440
;  PRIOR FILING DATE: 2000-07-20
;  PRIOR APPLICATION NUMBER: US 60/214,263
;  PRIOR FILING DATE: 2000-06-26
;  PRIOR APPLICATION NUMBER: US 60/173,754
;  PRIOR FILING DATE: 1999-12-30
;  NUMBER OF SEQ ID NOS: 409
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 129
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Conus striatus
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (1)..(30)
;   OTHER INFORMATION: Xaa at residues 3 and 5 may be Glu or gamma-carboxy-Glu; Xaa at r
;   OTHER INFORMATION: esidues 6, 7, 18 and 19 may be Pro or hydroxy-Pro; Xaa at residue
;   OTHER INFORMATION: s 24 and 28 may be Trp or bromo-Trp
US-09-749-637A-129

  Query Match             100.0%;  Score 33;  DB 3;  Length 30;
  Best Local Similarity   60.0%;  Pred. No. 9.5e+02;
  Matches    6;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|||:|:|
Db         14 FKIGXXCCSG 23


RESULT 44
US-09-749-637A-283
; Sequence 283, Application US/09749637A
; Patent No. US20020173449A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M.
;  APPLICANT:  Cartier, G. Edward
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Hillyard, David R.
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Layer, Richard T.
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: O-Superfamily Conotoxin Peptides
;  FILE REFERENCE: 2314-227
;  CURRENT APPLICATION NUMBER: US/09/749,637A
;  CURRENT FILING DATE:  2000-12-28
;  PRIOR APPLICATION NUMBER: US 60/243,412
;  PRIOR FILING DATE: 2000-10-27
;  PRIOR APPLICATION NUMBER: US60/219,440
;  PRIOR FILING DATE: 2000-07-20
;  PRIOR APPLICATION NUMBER: US 60/214,263
;  PRIOR FILING DATE: 2000-06-26
;  PRIOR APPLICATION NUMBER: US 60/173,754
;  PRIOR FILING DATE: 1999-12-30
;  NUMBER OF SEQ ID NOS: 409
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 283
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Conus geographus
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (1)..(30)
;   OTHER INFORMATION: Xaa at residues 3and 5 may be Glu or gamma-carboxy-Glu; Xaa at re
;   OTHER INFORMATION: sidues 6, 7, 18 and 19 may be Pro or hydroxy-Pro; Xaa at residues
;   OTHER INFORMATION: 24 and 28 may be Trp or bromo-Trp
US-09-749-637A-283

  Query Match             100.0%;  Score 33;  DB 3;  Length 30;
  Best Local Similarity   60.0%;  Pred. No. 9.5e+02;
  Matches    6;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|||:|:|
Db         14 FKIGXXCCSG 23


RESULT 45
US-10-433-485A-1
; Sequence 1, Application US/10433485A
; Publication No. US20040131617A1
; GENERAL INFORMATION:
;  APPLICANT: WHITE, Jay A.
;  APPLICANT:  PETKOVICH, P. Martin
;  APPLICANT:  JONES, Glenville
;  APPLICANT:  RAMSHAW, Heather
;  TITLE OF INVENTION: P450RAI-2(P450 Cytochrome 26B), Encoding Nucleic Acid
;  TITLE OF INVENTION:  Molecules and Methods and Uses Thereof
;  FILE REFERENCE: 11812-78
;  CURRENT APPLICATION NUMBER: US/10/433,485A
;  CURRENT FILING DATE:  2003-06-13
;  PRIOR APPLICATION NUMBER: PCT/CA01/01805
;  PRIOR FILING DATE: 2001-12-17
```

```
;  PRIOR APPLICATION NUMBER: PCT/CA00/01493
;  PRIOR FILING DATE: 2000-12-15
;  PRIOR APPLICATION NUMBER: 60/178,314
;  PRIOR FILING DATE: 2000-01-27
;  PRIOR APPLICATION NUMBER: 60/171,110
;  PRIOR FILING DATE: 1999-12-16
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: typical heme binding motif found in all Cytochrome P450s
US-10-433-485A-1

  Query Match             100.0%;  Score 33;  DB 4;  Length 30;
  Best Local Similarity   40.0%;  Pred. No. 9.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          8 FGIGKRVCMG 17


RESULT 46
US-10-839-227-126
; Sequence 126, Application US/10839227
; Publication No. US20050214903A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M.
;  APPLICANT:  Cartier, G. Edward
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Hillyard, David R.
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Layer, Richard T.
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: O-Superfamily Conotoxin Peptides
;  FILE REFERENCE: 2314-277
;  CURRENT APPLICATION NUMBER: US/10/839,227
;  CURRENT FILING DATE:  2004-05-06
;  PRIOR APPLICATION NUMBER: US 09/749,637
;  PRIOR FILING DATE: 2000-12-28
;  PRIOR APPLICATION NUMBER: US 60/243,412
;  PRIOR FILING DATE: 2000-10-27
;  PRIOR APPLICATION NUMBER: US60/219,440
;  PRIOR FILING DATE: 2000-07-20
;  PRIOR APPLICATION NUMBER: US 60/214,263
;  PRIOR FILING DATE: 2000-06-26
;  PRIOR APPLICATION NUMBER: US 60/173,754
;  PRIOR FILING DATE: 1999-12-30
;  NUMBER OF SEQ ID NOS: 409
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 126
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Conus marmoreus
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (1)..(30)
;   OTHER INFORMATION: Xaa at residues 3 and 5 may be Glu or gamma-carboxy-Glu; Xaa at r
;   OTHER INFORMATION: esidues 6, 7, 18 and 19 may be Pro or hydroxy-Pro; Xaa at residue
;   OTHER INFORMATION: s 24 and 28 may be Trp or bromo-Trp
US-10-839-227-126

  Query Match             100.0%;  Score 33;  DB 5;  Length 30;
  Best Local Similarity   60.0%;  Pred. No. 9.5e+02;
  Matches    6;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|||:|:|
Db         14 FKIGXXCCSG 23


RESULT 47
US-10-839-227-129
; Sequence 129, Application US/10839227
; Publication No. US20050214903A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M.
;  APPLICANT:  Cartier, G. Edward
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Hillyard, David R.
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Layer, Richard T.
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: O-Superfamily Conotoxin Peptides
;  FILE REFERENCE: 2314-277
;  CURRENT APPLICATION NUMBER: US/10/839,227
;  CURRENT FILING DATE:  2004-05-06
;  PRIOR APPLICATION NUMBER: US 09/749,637
;  PRIOR FILING DATE: 2000-12-28
;  PRIOR APPLICATION NUMBER: US 60/243,412
;  PRIOR FILING DATE: 2000-10-27
;  PRIOR APPLICATION NUMBER: US60/219,440
;  PRIOR FILING DATE: 2000-07-20
;  PRIOR APPLICATION NUMBER: US 60/214,263
;  PRIOR FILING DATE: 2000-06-26
;  PRIOR APPLICATION NUMBER: US 60/173,754
;  PRIOR FILING DATE: 1999-12-30
;  NUMBER OF SEQ ID NOS: 409
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 129
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Conus striatus
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (1)..(30)
;   OTHER INFORMATION: Xaa at residues 3 and 5 may be Glu or gamma-carboxy-Glu; Xaa at r
;   OTHER INFORMATION: esidues 6, 7, 18 and 19 may be Pro or hydroxy-Pro; Xaa at residue
;   OTHER INFORMATION: s 24 and 28 may be Trp or bromo-Trp
US-10-839-227-129

  Query Match             100.0%;  Score 33;  DB 5;  Length 30;
  Best Local Similarity   60.0%;  Pred. No. 9.5e+02;
  Matches    6;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|||:|:|
Db         14 FKIGXXCCSG 23


RESULT 48
US-10-839-227-283
; Sequence 283, Application US/10839227
; Publication No. US20050214903A1
; GENERAL INFORMATION:
;  APPLICANT: University of Utah Research Foundation
;  APPLICANT:  Cognetix, Inc.
;  APPLICANT:  Olivera, Baldomero M.
;  APPLICANT:  Cartier, G. Edward
;  APPLICANT:  Watkins, Maren
;  APPLICANT:  Hillyard, David R.
;  APPLICANT:  McIntosh, J. Michael
;  APPLICANT:  Layer, Richard T.
;  APPLICANT:  Jones, Robert M.
;  TITLE OF INVENTION: O-Superfamily Conotoxin Peptides
;  FILE REFERENCE: 2314-277
;  CURRENT APPLICATION NUMBER: US/10/839,227
;  CURRENT FILING DATE:  2004-05-06
;  PRIOR APPLICATION NUMBER: US 09/749,637
```

```
;  PRIOR FILING DATE: 2000-12-28
;  PRIOR APPLICATION NUMBER: US 60/243,412
;  PRIOR FILING DATE: 2000-10-27
;  PRIOR APPLICATION NUMBER: US60/219,440
;  PRIOR FILING DATE: 2000-07-20
;  PRIOR APPLICATION NUMBER: US 60/214,263
;  PRIOR FILING DATE: 2000-06-26
;  PRIOR APPLICATION NUMBER: US 60/173,754
;  PRIOR FILING DATE: 1999-12-30
;  NUMBER OF SEQ ID NOS: 409
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 283
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Conus geographus
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (1)..(30)
;   OTHER INFORMATION: Xaa at residues 3and 5 may be Glu or gamma-carboxy-Glu; Xaa at re
;   OTHER INFORMATION: sidues 6, 7, 18 and 19 may be Pro or hydroxy-Pro; Xaa at residues
;   OTHER INFORMATION: 24 and 28 may be Trp or bromo-Trp
US-10-839-227-283

  Query Match             100.0%;  Score 33;  DB 5;  Length 30;
  Best Local Similarity   60.0%;  Pred. No. 9.5e+02;
  Matches    6;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|||:|:|
Db         14 FKIGXXCCSG 23


RESULT 49
US-10-437-963-159510
; Sequence 159510, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 159510
;   LENGTH: 32
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_5887C.1.pep
US-10-437-963-159510

  Query Match             100.0%;  Score 33;  DB 4;  Length 32;
  Best Local Similarity   40.0%;  Pred. No. 1e+03;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              .|::|:::|:|
Db          5 FICGIWSCEG 14


RESULT 50
US-10-351-641-487
; Sequence 487, Application US/10351641
; Publication No. US20030186874A1
; GENERAL INFORMATION:
;  APPLICANT: Barney, S.
;  APPLICANT:  Guthrie, K.
;  APPLICANT:  Merutka, G.
;  APPLICANT:  Anwer, M.
;  APPLICANT:  Lambert, D.
;  TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
;  TITLE OF INVENTION:  PROPERTIES
;  FILE REFERENCE: 7872-100
;  CURRENT APPLICATION NUMBER: US/10/351,641
;  CURRENT FILING DATE:  2003-01-24
;  PRIOR APPLICATION NUMBER: 09/350,641
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: 09/315,304
;  PRIOR FILING DATE: 1999-05-20
;  PRIOR APPLICATION NUMBER: 09/082,279
;  PRIOR FILING DATE: 1998-05-20
;  NUMBER OF SEQ ID NOS: 1757
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 487
;   LENGTH: 35
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Core polypeptide
US-10-351-641-487

  Query Match             100.0%;  Score 33;  DB 4;  Length 35;
  Best Local Similarity   40.0%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         26 FLGGTTVCLG 35


Search completed: March  8, 2006, 11:28:19
Job time : 167 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2006, 11:16:36 ; Search time 189 Seconds
(without alignments)
23.248 Million cell updates/sec

Title: US-10-751-235-14
Perfect score: 33
Sequence: 1 FXXGXXXCXG 10

Scoring table: BLOSUM62DX* – this matrix allows an Xaa in the query seq to match anything in the database sequence.
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163 These matches are counted as conservative substitutions in the alignment & are marked with a colon (:).

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

50 alignments saved as requested

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	100.0	10	2	AAY09193	Aay09193 Soybean c
2	33	100.0	10	3	AAY91408	Aay91408 Human sec
3	33	100.0	10	4	AAM10370	Aam10370 HLA-B *07
4	33	100.0	10	4	AAM13360	Aam13360 HLA-A26 d
5	33	100.0	10	4	AAM09243	Aam09243 HLA-A1 de
6	33	100.0	10	4	AAM10626	Aam10626 HLA-A *02
7	33	100.0	10	5	ABP53084	Abp53084 Conserved
8	33	100.0	10	5	ABP53085	Abp53085 CYP 27C1
9	33	100.0	10	5	AAE25791	Aae25791 Aspergill
10	33	100.0	10	6	ABU08352	Abu08352 Conserved
11	33	100.0	10	6	ABR42010	Abr42010 Cytochrom
12	33	100.0	10	6	AAE35263	Aae35263 Cytochrom
13	33	100.0	10	6	AAE35262	Aae35262 Cytochrom
14	33	100.0	10	7	ADC56716	Adc56716 Peptide 1
15	33	100.0	10	7	ADC98703	Adc98703 Cytochrom
16	33	100.0	10	9	AEB16917	Aeb16917 LUT1 cyto
17	33	100.0	10	9	AEB16916	Aeb16916 Cytochrom
18	33	100.0	11	2	AAR30867	Aar30867 Consensus
19	33	100.0	11	2	AAR40877	Aar40877 SSP for f
20	33	100.0	11	2	AAR40874	Aar40874 SSP for f
21	33	100.0	11	2	AAR40876	Aar40876 SSP for f
22	33	100.0	11	2	AAR41633	Aar41633 SSP for f
23	33	100.0	11	2	AAR40872	Aar40872 SSP for f
24	33	100.0	11	2	AAR40879	Aar40879 SSP for f
25	33	100.0	11	2	AAR41632	Aar41632 SSP for f
26	33	100.0	11	2	AAR41634	Aar41634 SSP for f
27	33	100.0	11	2	AAR40873	Aar40873 SSP for f
28	33	100.0	11	2	AAR40878	Aar40878 SSP for f
29	33	100.0	11	2	AAR40880	Aar40880 SSP for f
30	33	100.0	11	2	AAR40870	Aar40870 SSP for f
31	33	100.0	11	2	AAR40871	Aar40871 SSP for f
32	33	100.0	11	2	AAR41635	Aar41635 SSP for f
33	33	100.0	11	2	AAR40875	Aar40875 SSP for f
34	33	100.0	11	5	AAU80018	Aau80018 Conserved
35	33	100.0	11	5	ADO58606	Ado58606 Pepper cy
36	33	100.0	11	6	AAE30014	Aae30014 Cytochrom
37	33	100.0	13	2	AAR30868	Aar30868 Consensus
38	33	100.0	13	2	AAY09192	Aay09192 Soybean c
39	33	100.0	13	3	AAB11394	Aab11394 Gerbera f
40	33	100.0	14	4	AAG98132	Aag98132 Human SNP
41	33	100.0	15	1	AAP20076	Aap20076 Synthetic
42	33	100.0	15	1	AAP50146	Aap50146 Sequence
43	33	100.0	15	1	AAP60758	Aap60758 Labeled s
44	33	100.0	15	2	AAR24422	Aar24422 Sequence
45	33	100.0	15	2	AAR82090	Aar82090 Hepatitis
46	33	100.0	15	3	AAB19305	Aab19305 Immunogen
47	33	100.0	15	5	AAU91319	Aau91319 Cytochrom
48	33	100.0	15	6	ABG75617	Abg75617 Arabidops
49	33	100.0	15	9	ADV23624	Adv23624 HBV immun
50	33	100.0	15	9	ADV23625	Adv23625 HBV immun
51	33	100.0	15	9	ADV23253	Adv23253 HBV immun
52	33	100.0	16	2	AAR38307	Aar38307 N-termina
53	33	100.0	16	3	AAY65096	Aay65096 Human 5'
54	33	100.0	16	6	ABJ26710	Abj26710 Seed deve
55	33	100.0	16	8	ADU72660	Adu72660 Signal pe
56	33	100.0	16	9	ADZ73651	Adz73651 Human inc
57	33	100.0	18	2	AAW15123	Aaw15123 Peptide e
58	33	100.0	20	2	AAY13668	Aay13668 Erythropo
59	33	100.0	20	2	AAW26982	Aaw26982 Monomer s
60	33	100.0	20	4	AAU05333	Aau05333 R1 and R2
61	33	100.0	20	5	ABJ09980	Abj09980 Hepatitis
62	33	100.0	20	5	AAE15327	Aae15327 Zebrafish
63	33	100.0	20	7	ADW34802	Adw34802 HLA bindi
64	33	100.0	20	7	ADW33569	Adw33569 HLA bindi
65	33	100.0	20	7	ADW36246	Adw36246 HLA bindi
66	33	100.0	20	8	ADU66830	Adu66830 CYP4503A1
67	33	100.0	20	8	ADU66828	Adu66828 RATCYP4A1
68	33	100.0	20	8	ADU66829	Adu66829 RABCYP4A5
69	33	100.0	20	8	ADU66831	Adu66831 hCYTFAOH
70	33	100.0	20	8	ADU66827	Adu66827 ATCYTP450
71	33	100.0	20	9	ADV90770	Adv90770 CYP4503A1
72	33	100.0	20	9	ADV90768	Adv90768 RATCYP4A1
73	33	100.0	20	9	ADV90767	Adv90767 ATCYTP450
74	33	100.0	20	9	ADV90771	Adv90771 hCYTFAOH
75	33	100.0	20	9	ADV90769	Adv90769 RABCYP4A5
76	33	100.0	20	9	ADY62584	Ady62584 Cytochrom
77	33	100.0	20	9	ADY62585	Ady62585 Cytochrom
78	33	100.0	20	9	ADY62587	Ady62587 Cytochrom
79	33	100.0	20	9	ADY62588	Ady62588 Cytochrom
80	33	100.0	20	9	ADY62586	Ady62586 Cytochrom
81	33	100.0	21	5	AAE25786	Aae25786 Aspergill
82	33	100.0	22	4	AAB87046	Aab87046 Human TAN
83	33	100.0	23	4	AAB87038	Aab87038 Human TAN
84	33	100.0	23	4	AAB49519	Aab49519 Clone HCE
85	33	100.0	30	4	AAU05872	Aau05872 Cone snai
86	33	100.0	30	4	AAU05874	Aau05874 Cone snai
87	33	100.0	30	4	AAU05978	Aau05978 Cone snai
88	33	100.0	30	4	AAB85150	Aab85150 Cytochrom
89	33	100.0	30	5	ABP52141	Abp52141 Cytochrom
90	33	100.0	33	8	ADL97623	Adl97623 Protein e
91	33	100.0	34	8	ADM78246	Adm78246 Hepatitis
92	33	100.0	34	8	ADM78312	Adm78312 Hepatitis
93	33	100.0	34	8	ADM78221	Adm78221 Hepatitis
94	33	100.0	34	8	ADM78254	Adm78254 Hepatitis
95	33	100.0	35	3	AAY89167	Aay89167 Core poly
96	33	100.0	35	3	AAY89126	Aay89126 Core poly
97	33	100.0	35	3	AAY89164	Aay89164 Core poly

98	33	100.0	35	3	AAY89129	Aay89129 Core poly
99	33	100.0	35	3	AAY89127	Aay89127 Core poly
100	33	100.0	35	3	AAY89163	Aay89163 Core poly

ALIGNMENTS

RESULT 1
AAY09193
ID AAY09193 standard; peptide; 10 AA.
XX
AC AAY09193;
XX
DT 20-JUL-1999 (first entry)
XX
DE Soybean cytochrome P450 enzyme conserved peptide motif.
XX
KW Soybean; Solanaceae crop plant; cytochrome P450; transgenic plant;
KW enzyme; phenylurea herbicide; herbicide resistance.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
FT Misc-difference 1. .10
FT /note= "Xaa is any amino acid"
XX
PN WO9919493-A2.
XX
PD 22-APR-1999.
XX
PF 05-OCT-1998; 98WO-US020807.
XX
PR 10-OCT-1997; 97US-00948564.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Siminszky B, Dewey RE, Corbin FT;
XX
DR WPI; 1999-302532/25.
XX
PT DNA encoding soybean cytochrome P450 enzymes.
XX
PS Example 2; Page 26; 93pp; English.
XX
CC The invention provides new DNA molecules (AAX60773-X60781) encoding
CC soybean cytochrome P450 enzymes (AAY09183-Y09191) respectively. The DNA
CC encoding the cytochrome P450 enzymes is useful for transformation of
CC Solanaceae crop plants. Transgenic plants comprising DNA constructs
CC having the P450 encoding nucleic acid sequences are resistant to
CC phenylurea herbicides. The transgenic plants have increased resistance to
CC phenylurea herbicides compared to wild-type plants of the same species.
CC The plant crops, e.g. turfgrass, tobacco, potato, tomato, corn, rice,
CC cotton, soybean, rape, wheat, oats, barley or rice are particularly
CC resistant to fluometuron, linuron, chlortoluron or diuron
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 33; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
||||||||||
Db 1 FXXGXXXCXG 10

RESULT 2
AAY91408
ID AAY91408 standard; protein; 10 AA.
XX
AC AAY91408;
XX
DT 29-JUN-2000 (first entry)
XX
DE Human secreted protein sequence encoded by gene 3 SEQ ID NO:129.
XX
KW Human; secreted protein; diagnosis; neuroprotective; nootropic;
KW neuroleptic; antimanic; cerebroprotective; immunomodulatory;
KW anti-microbial; cardiant; cytostatic; antiinflammatory; haemostatic;
KW anticonvulsant; vasotropic; vaccine; gene therapy; anti-sense therapy;
KW neural; reproductive; immune disorder; immunodeficiency; infection;
KW lymphoma; demyelinating disease; autoimmunity; cancer; inflammation;
KW aneurysm; haemorrhage; Alzheimer's disease; Parkinson's disease;
KW Huntington's disease; Tourette syndrome; multiple sclerosis; meningitis;
KW ischaemia; mania; dementia; obsessive compulsive disorder;
KW viral prophylaxis; developmental disorder; sexually-linked disorder;
KW cardiovascular disorder; food additive; preservative; chromosome 11.
XX
OS Homo sapiens.
XX
PN WO200011014-A1.
XX
PD 02-MAR-2000.
XX
PF 24-AUG-1999; 99WO-US019330.
XX
PR 25-AUG-1998; 98US-0097917P.
PR 31-AUG-1998; 98US-0098634P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Moore PA, Ruben SM, Olsen HS, Shi Y, Rosen CA, Florence KA;
PI Soppet DR, Lafleur DW, Endress GA, Ebner R, Komatsoulis G, Duan RD;
XX
DR WPI; 2000-224656/19.
XX
PT Novel secreted proteins and corresponding DNA molecules that can be used
PT to prevent, treat and diagnose disease in humans, for example,
PT Alzheimer's, cancer, and immune disorders.
XX
PS Disclosure; Page 390; 416pp; English.
XX
CC The polynucleotide sequences given in AAA26281 to AAA26336 encode the
CC human secreted proteins given in AAY91346 to AAY91449. The human secreted
CC proteins can have activities based on the tissues and cells they are
CC expressed in. Examples of the activities are: neuroprotective; nootropic;
CC neuroleptic; antimanic; cerebroprotective; immunomodulatory; anti-
CC microbial; cardiant; cytostatic; antiinflammatory; haemostatic;
CC anticonvulsant; and vasotropic. The polynucleotides and proteins may be
CC used to prevent, treat or ameliorate a medical condition, e.g. by protein
CC or gene therapy. Conditions treatable by the proteins of the invention
CC include neural, reproductive, or immune disorders, especially
CC immunodeficiency, infection, lymphomas, demyelinating diseases, auto-
CC immunities, cancer, general microbial infection, inflammation, aneurysms
CC and haemorrhages. Specific examples include: Alzheimer's disease;
CC Parkinson's; Huntington's; Tourette syndrome; multiple sclerosis;
CC meningitis; ischaemia; prostate cancer; mania; dementia; obsessive
CC compulsive disorder and viral prophylaxis. The polynucleotides and
CC proteins can also be used in the detection of disorders associated with
CC the function of the protein, for example, the detection of developmental
CC disorders, sexually-linked disorders, or disorders of the cardiovascular
CC system. They may also be used as food additives or preservatives.
CC AAA26272 to AAA26280 and AAY91345 are sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 33; DB 3; Length 10;
Best Local Similarity 40.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|::|:::|:|
Db 1 FSLGRRHCLG 10

```
RESULT 3
AAM10370
ID   AAM10370 standard; peptide; 10 AA.
XX
AC   AAM10370;
XX
DT   09-OCT-2001  (first entry)
XX
DE   HLA-B *0702 decamer #320.
XX
KW   Cytochrome P450 1B1; CYP1B1; cancer; immunotherapy; diagnosis; HLA;
KW   human leukocyte antigen; cytotoxic T-lymphocyte; CTL; MHC; APC;
KW   major histocompatibility complex; antigen presenting cell; melanoma;
KW   lymphoma; carcinoma; sarcoma; multiple myeloma; leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200135810-A2.
XX
PD   25-MAY-2001.
XX
PF   15-NOV-2000; 2000WO-US031513.
XX
PR   15-NOV-1999;   99US-0165590P.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
PA   (UYBO-) UNIV BOSTON.
XX
PI   Schultze JL,  Vonderheide RH,  Sherr D,  Nadler LM,  Maecker B;
PI   Von Bergwelt- Baildon M;
XX
DR   WPI; 2001-355537/37.
XX
PT   Treating a cancer patient involves administering to patient a cytotoxic T
PT   -lymphocyte, an antigen presenting cell that activates T cells, a peptide
PT   of cytochrome P450 1B1, or a polynucleotide encoding the peptide.
XX
PS   Disclosure; Page 76; 121pp; English.
XX
CC   The present invention describes a method for treating a patient having,
CC   or is at risk of having, a cell expressing cytochrome P450 1B1 (CYP1B1),
CC   which involves administering to the patient a cytotoxic T-lymphocyte
CC   (CTL) that kills the cell, an antigen presenting cell (APC) that
CC   activates CTL, a peptide of CYP1B1 that binds to major histocompatibility
CC   complex (MHC) molecule, or a nucleic acid molecule encoding CYP1B1 or
CC   peptide of CYP1B1. Also described are: (1) a method for assessing the
CC   level of immunity of a patient to CYP1B1 or a peptide of CYP1B1 that
CC   binds to MHC complex molecule, involving measuring the level of CTL
CC   specific for CYP1B1 or the peptide of CYP1B1 in a sample from the patient
CC   ; (2) a CYP1B1 peptide (I) that binds to MHC complex molecule; (3) an ex
CC   vivo generated CTL (II) that specifically kills a cell expressing CYP1B1
CC   in a specific of a MHC complex-restricted fashion; and (4) an ex vivo
CC   generated APC (III) that presents a peptide of a CYP1B1 in the context of
CC   a MHC complex molecule. The method is useful for treating a patient
CC   having or is at risk of having a cell that expresses CYP1B1. The method
CC   is useful for the prevention, treatment and diagnosis of cancer, e.g.
CC   melanoma, lymphoma, carcinoma, sarcoma, multiple myeloma, leukemia, and
CC   lung, ovarian, uterine, cervical, prostate, liver, colon, pancreatic and
CC   brain cancer. AAM06905 to AAM13566 represent CYP1B1 peptides which can
CC   bind to human leukocyte antigens (HLAs), as well as other amino acid
CC   sequence used in the exemplification of the present invention
XX
SQ   Sequence 10 AA;

  Query Match           100.0%;  Score 33;  DB 4;  Length 10;
  Best Local Similarity  40.0%;  Pred. No. 3.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FSVGKRRCIG 10
```

```
RESULT 4
AAM13360
ID   AAM13360 standard; peptide; 10 AA.
XX
AC   AAM13360;
XX
DT   09-OCT-2001  (first entry)
XX
DE   HLA-A26 decamer #423.
XX
KW   Cytochrome P450 1B1; CYP1B1; cancer; immunotherapy; diagnosis; HLA;
KW   human leukocyte antigen; cytotoxic T-lymphocyte; CTL; MHC; APC;
KW   major histocompatibility complex; antigen presenting cell; melanoma;
KW   lymphoma; carcinoma; sarcoma; multiple myeloma; leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200135810-A2.
XX
PD   25-MAY-2001.
XX
PF   15-NOV-2000; 2000WO-US031513.
XX
PR   15-NOV-1999;   99US-0165590P.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
PA   (UYBO-) UNIV BOSTON.
XX
PI   Schultze JL,  Vonderheide RH,  Sherr D,  Nadler LM,  Maecker B;
PI   Von Bergwelt- Baildon M;
XX
DR   WPI; 2001-355537/37.
XX
PT   Treating a cancer patient involves administering to patient a cytotoxic T
PT   -lymphocyte, an antigen presenting cell that activates T cells, a peptide
PT   of cytochrome P450 1B1, or a polynucleotide encoding the peptide.
XX
PS   Disclosure; Page 95; 121pp; English.
XX
CC   The present invention describes a method for treating a patient having,
CC   or is at risk of having, a cell expressing cytochrome P450 1B1 (CYP1B1),
CC   which involves administering to the patient a cytotoxic T-lymphocyte
CC   (CTL) that kills the cell, an antigen presenting cell (APC) that
CC   activates CTL, a peptide of CYP1B1 that binds to major histocompatibility
CC   complex (MHC) molecule, or a nucleic acid molecule encoding CYP1B1 or
CC   peptide of CYP1B1. Also described are: (1) a method for assessing the
CC   level of immunity of a patient to CYP1B1 or a peptide of CYP1B1 that
CC   binds to MHC complex molecule, involving measuring the level of CTL
CC   specific for CYP1B1 or the peptide of CYP1B1 in a sample from the patient
CC   ; (2) a CYP1B1 peptide (I) that binds to MHC complex molecule; (3) an ex
CC   vivo generated CTL (II) that specifically kills a cell expressing CYP1B1
CC   in a specific of a MHC complex-restricted fashion; and (4) an ex vivo
CC   generated APC (III) that presents a peptide of a CYP1B1 in the context of
CC   a MHC complex molecule. The method is useful for treating a patient
CC   having or is at risk of having a cell that expresses CYP1B1. The method
CC   is useful for the prevention, treatment and diagnosis of cancer, e.g.
CC   melanoma, lymphoma, carcinoma, sarcoma, multiple myeloma, leukemia, and
CC   lung, ovarian, uterine, cervical, prostate, liver, colon, pancreatic and
CC   brain cancer. AAM06905 to AAM13566 represent CYP1B1 peptides which can
CC   bind to human leukocyte antigens (HLAs), as well as other amino acid
CC   sequence used in the exemplification of the present invention
XX
SQ   Sequence 10 AA;

  Query Match           100.0%;  Score 33;  DB 4;  Length 10;
  Best Local Similarity  40.0%;  Pred. No. 3.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FSVGKRRCIG 10
```

```
RESULT 5
AAM09243
ID   AAM09243 standard; peptide; 10 AA.
XX
AC   AAM09243;
XX
DT   09-OCT-2001  (first entry)
XX
DE   HLA-A1 decamer #110.
XX
KW   Cytochrome P450 1B1; CYP1B1; cancer; immunotherapy; diagnosis; HLA;
KW   human leukocyte antigen; cytotoxic T-lymphocyte; CTL; MHC; APC;
KW   major histocompatibility complex; antigen presenting cell; melanoma;
KW   lymphoma; carcinoma; sarcoma; multiple myeloma; leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200135810-A2.
XX
PD   25-MAY-2001.
XX
PF   15-NOV-2000; 2000WO-US031513.
XX
PR   15-NOV-1999;   99US-0165590P.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
PA   (UYBO-) UNIV BOSTON.
XX
PI   Schultze JL,  Vonderheide RH,  Sherr D,  Nadler LM,  Maecker B;
PI   Von Bergwelt- Baildon M;
XX
DR   WPI; 2001-355537/37.
XX
PT   Treating a cancer patient involves administering to patient a cytotoxic T
PT   -lymphocyte, an antigen presenting cell that activates T cells, a peptide
PT   of cytochrome P450 1B1, or a polynucleotide encoding the peptide.
XX
PS   Disclosure; Page 70; 121pp; English.
XX
CC   The present invention describes a method for treating a patient having,
CC   or is at risk of having, a cell expressing cytochrome P450 1B1 (CYP1B1),
CC   which involves administering to the patient a cytotoxic T-lymphocyte
CC   (CTL) that kills the cell, an antigen presenting cell (APC) that
CC   activates CTL, a peptide of CYP1B1 that binds to major histocompatibility
CC   complex (MHC) molecule, or a nucleic acid molecule encoding CYP1B1 or
CC   peptide of CYP1B1. Also described are: (1) a method for assessing the
CC   level of immunity of a patient to CYP1B1 or a peptide of CYP1B1 that
CC   binds to MHC complex molecule, involving measuring the level of CTL
CC   specific for CYP1B1 or the peptide of CYP1B1 in a sample from the patient
CC   ; (2) a CYP1B1 peptide (I) that binds to MHC complex molecule; (3) an ex
CC   vivo generated CTL (II) that specifically kills a cell expressing CYP1B1
CC   in a specific of a MHC complex-restricted fashion; and (4) an ex vivo
CC   generated APC (III) that presents a peptide of a CYP1B1 in the context of
CC   a MHC complex molecule. The method is useful for treating a patient
CC   having or is at risk of having a cell that expresses CYP1B1. The method
CC   is useful for the prevention, treatment and diagnosis of cancer, e.g.
CC   melanoma, lymphoma, carcinoma, sarcoma, multiple myeloma, leukemia, and
CC   lung, ovarian, uterine, cervical, prostate, liver, colon, pancreatic and
CC   brain cancer. AAM06905 to AAM13566 represent CYP1B1 peptides which can
CC   bind to human leukocyte antigens (HLAs), as well as other amino acid
CC   sequence used in the exemplification of the present invention
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 33;  DB 4;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 3.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FSVGKRRCIG 10
```

```
RESULT 6
AAM10626
ID   AAM10626 standard; peptide; 10 AA.
XX
AC   AAM10626;
XX
DT   09-OCT-2001  (first entry)
XX
DE   HLA-A *0201 decamer #409.
XX
KW   Cytochrome P450 1B1; CYP1B1; cancer; immunotherapy; diagnosis; HLA;
KW   human leukocyte antigen; cytotoxic T-lymphocyte; CTL; MHC; APC;
KW   major histocompatibility complex; antigen presenting cell; melanoma;
KW   lymphoma; carcinoma; sarcoma; multiple myeloma; leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200135810-A2.
XX
PD   25-MAY-2001.
XX
PF   15-NOV-2000; 2000WO-US031513.
XX
PR   15-NOV-1999;   99US-0165590P.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
PA   (UYBO-) UNIV BOSTON.
XX
PI   Schultze JL,  Vonderheide RH,  Sherr D,  Nadler LM,  Maecker B;
PI   Von Bergwelt- Baildon M;
XX
DR   WPI; 2001-355537/37.
XX
PT   Treating a cancer patient involves administering to patient a cytotoxic T
PT   -lymphocyte, an antigen presenting cell that activates T cells, a peptide
PT   of cytochrome P450 1B1, or a polynucleotide encoding the peptide.
XX
PS   Disclosure; Page 78; 121pp; English.
XX
CC   The present invention describes a method for treating a patient having,
CC   or is at risk of having, a cell expressing cytochrome P450 1B1 (CYP1B1),
CC   which involves administering to the patient a cytotoxic T-lymphocyte
CC   (CTL) that kills the cell, an antigen presenting cell (APC) that
CC   activates CTL, a peptide of CYP1B1 that binds to major histocompatibility
CC   complex (MHC) molecule, or a nucleic acid molecule encoding CYP1B1 or
CC   peptide of CYP1B1. Also described are: (1) a method for assessing the
CC   level of immunity of a patient to CYP1B1 or a peptide of CYP1B1 that
CC   binds to MHC complex molecule, involving measuring the level of CTL
CC   specific for CYP1B1 or the peptide of CYP1B1 in a sample from the patient
CC   ; (2) a CYP1B1 peptide (I) that binds to MHC complex molecule; (3) an ex
CC   vivo generated CTL (II) that specifically kills a cell expressing CYP1B1
CC   in a specific of a MHC complex-restricted fashion; and (4) an ex vivo
CC   generated APC (III) that presents a peptide of a CYP1B1 in the context of
CC   a MHC complex molecule. The method is useful for treating a patient
CC   having or is at risk of having a cell that expresses CYP1B1. The method
CC   is useful for the prevention, treatment and diagnosis of cancer, e.g.
CC   melanoma, lymphoma, carcinoma, sarcoma, multiple myeloma, leukemia, and
CC   lung, ovarian, uterine, cervical, prostate, liver, colon, pancreatic and
CC   brain cancer. AAM06905 to AAM13566 represent CYP1B1 peptides which can
CC   bind to human leukocyte antigens (HLAs), as well as other amino acid
CC   sequence used in the exemplification of the present invention
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 33;  DB 4;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 3.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FSVGKRRCIG 10
```

```
RESULT 7
ABP53084
ID   ABP53084 standard; peptide; 10 AA.
XX
AC   ABP53084;
XX
DT   11-NOV-2002  (first entry)
XX
DE   Conserved haem region generic peptide.
XX
KW   Cloning; characterisation; human; cytochrome P450; CYP 27C1; cytostatic;
KW   thyromimetic; antidiabetic; antipsoriatic; tuberculostatic; osteopathic;
KW   dermatological; antilipaemic; gene therapy; vaccine; Vitamin D; diabetes;
KW   vitamin D metabolite deficiency; hyperparathyroidism; hypoparathyroidism;
KW   medullary carcinoma; psoriasis; sarcoidosis; tuberculosis; osteomalacia;
KW   chronic renal disease; vitamin D dependent rickets; anticonvulsant;
KW   fibrogenesis imperfecta ossium; osteititis fibrosa cystica; osteoporosis;
KW   osteopaenia; osteosclerosis; renal osteodystrophy; rickets; steatorrhoea;
KW   glucocorticoid antagonism; idiopathic hypercalcaemia; tropical sprue;
KW   malabsorption syndrome; cholesterol steroid; lipid metabolic disorder.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 2
FT                   /note= "unspecified"
FT   Misc-difference 3
FT                   /note= "unspecified"
FT   Misc-difference 5
FT                   /note= "unspecified"
FT   Misc-difference 6
FT                   /note= "unspecified"
FT   Misc-difference 7
FT                   /note= "unspecified"
FT   Misc-difference 9
FT                   /note= "unspecified"
XX
PN   WO200264765-A2.
XX
PD   22-AUG-2002.
XX
PF   11-FEB-2002; 2002WO-CA000163.
XX
PR   09-FEB-2001; 2001US-0267410P.
XX
PA   (CYTO-) CYTOCHROMA INC.
XX
PI   Wisniewski J;
XX
DR   WPI; 2002-657595/70.
XX
PT   New nucleic acid molecules encoding cytochrome P450 proteins, human CYP
PT   27C1 and a hybrid homologs from Xenopus laevis, useful for treating
PT   diseases related to vitamin D or vitamin D metabolite deficiency, e.g.
PT   parathyroidism and diabetes.
XX
PS   Example 1; Page 86; 209pp; English.
XX
CC   The present invention describes an isolated nucleic acid molecule (I)
CC   encoding human cytochrome P450, CYP 27C1, and a hybrid homologue from
CC   Xenopus laevis. (I) has thyromimetic, antidiabetic, cytostatic,
CC   antipsoriatic, tuberculostatic, osteopathic, dermatological and
CC   antilipaemic activities, and can be used in gene therapy and in vaccines.
CC   The nucleic acid molecules, proteins and methods from the present
CC   invention are useful for treating diseases related to vitamin D or
CC   vitamin D metabolite deficiency, e.g. hyper- and hypo-parathyroidism,
CC   Osudohypo-parathyroidism, Secondary hyperparathyroidism, diabetes,
CC   medullary carcinoma, psoriasis, sarcoidosis, tuberculosis, chronic renal
CC   disease, hypophosphatemic VDRR, vitamin D dependent rickets,
CC   anticonvulsant treatment, fibrogenesis imperfecta ossium, osteititis
CC   fibrosa cystica, osteomalacia, osteoporosis, osteopaenia, osteosclerosis,
CC   renal osteodystrophy, rickets, glucocorticoid antagonism, idiopathic
CC   hypercalcaemia, malabsorption syndrome, steatorrhoea, and tropical sprue,
CC   or cholesterol, steroid and other lipid metabolic disorders. The present
CC   sequence represents a conserved haem region generic peptide, which is
CC   used in an example from the present invention
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 33;  DB 5;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              ||||||||||
Db          1 FXXGXXXCXG 10


RESULT 8
ABP53085
ID   ABP53085 standard; peptide; 10 AA.
XX
AC   ABP53085;
XX
DT   11-NOV-2002  (first entry)
XX
DE   CYP 27C1 haem-binding signature peptide SEQ ID NO:47.
XX
KW   Cloning; characterisation; human; cytochrome P450; CYP 27C1; cytostatic;
KW   thyromimetic; antidiabetic; antipsoriatic; tuberculostatic; osteopathic;
KW   dermatological; antilipaemic; gene therapy; vaccine; Vitamin D; diabetes;
KW   vitamin D metabolite deficiency; hyperparathyroidism; hypoparathyroidism;
KW   medullary carcinoma; psoriasis; sarcoidosis; tuberculosis; osteomalacia;
KW   chronic renal disease; vitamin D dependent rickets; anticonvulsant;
KW   fibrogenesis imperfecta ossium; osteititis fibrosa cystica; osteoporosis;
KW   osteopaenia; osteosclerosis; renal osteodystrophy; rickets; steatorrhoea;
KW   glucocorticoid antagonism; idiopathic hypercalcaemia; tropical sprue;
KW   malabsorption syndrome; cholesterol steroid; lipid metabolic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200264765-A2.
XX
PD   22-AUG-2002.
XX
PF   11-FEB-2002; 2002WO-CA000163.
XX
PR   09-FEB-2001; 2001US-0267410P.
XX
PA   (CYTO-) CYTOCHROMA INC.
XX
PI   Wisniewski J;
XX
DR   WPI; 2002-657595/70.
XX
PT   New nucleic acid molecules encoding cytochrome P450 proteins, human CYP
PT   27C1 and a hybrid homologs from Xenopus laevis, useful for treating
PT   diseases related to vitamin D or vitamin D metabolite deficiency, e.g.
PT   parathyroidism and diabetes.
XX
PS   Example 1; Page 16; 209pp; English.
XX
CC   The present invention describes an isolated nucleic acid molecule (I)
CC   encoding human cytochrome P450, CYP 27C1, and a hybrid homologue from
CC   Xenopus laevis. (I) has thyromimetic, antidiabetic, cytostatic,
CC   antipsoriatic, tuberculostatic, osteopathic, dermatological and
CC   antilipaemic activities, and can be used in gene therapy and in vaccines.
CC   The nucleic acid molecules, proteins and methods from the present
CC   invention are useful for treating diseases related to vitamin D or
CC   vitamin D metabolite deficiency, e.g. hyper- and hypo-parathyroidism,
CC   Osudohypo-parathyroidism, Secondary hyperparathyroidism, diabetes,
CC   medullary carcinoma, psoriasis, sarcoidosis, tuberculosis, chronic renal
CC   disease, hypophosphatemic VDRR, vitamin D dependent rickets,
CC   anticonvulsant treatment, fibrogenesis imperfecta ossium, osteititis
CC   fibrosa cystica, osteomalacia, osteoporosis, osteopaenia, osteosclerosis,
```

```
CC   renal osteodystrophy, rickets, glucocorticoid antagonism, idiopathic
CC   hypercalcaemia, malabsorption syndrome, steatorrhoea, and tropical sprue,
CC   or cholesterol, steroid and other lipid metabolic disorders. The present
CC   sequence represents a CYP 27C1 haem-binding signature peptide, which is
CC   used in an example from the present invention
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 33;  DB 5;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 3.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGHGVRSCIG 10


RESULT 9
AAE25791
ID   AAE25791 standard; peptide; 10 AA.
XX
AC   AAE25791;
XX
DT   04-NOV-2002  (first entry)
XX
DE   Aspergillus ochraceus cytochrome p450 conserved haem binding motif.
XX
KW   11 alpha hydroxylase; enzyme; sitosterol; eplerenone; cell therapy;
KW   steroid bioconversion; antiinflammatory; antiarthritic; cytostatic;
KW   cardiant; cytochrome P450; oxidoreductase; haem binding motif.
XX
OS   Aspergillus ochraceus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 2. .3
FT                   /label= Unknown
FT   Misc-difference 5. .7
FT                   /label= Unknown
FT   Misc-difference 9
FT                   /label= Unknown
XX
PN   WO200246386-A2.
XX
PD   13-JUN-2002.
XX
PF   26-OCT-2001; 2001WO-US051070.
XX
PR   30-OCT-2000; 2000US-0244300P.
XX
PA   (PHAA ) PHARMACIA CORP.
PA   (BOLT/) BOLTON S.
PA   (CLAY/) CLAYTON R.
PA   (EAST/) EASTON A.
PA   (ENGE/) ENGEL L.
PA   (MESS/) MESSING D.
XX
PI   Bolton S,  Clayton R,  Easton A,  Engel L,  Messing D;
XX
DR   WPI; 2002-547772/58.
XX
PT   New isolated Aspergillus ochraceus 11 alpha-hydroxylase or
PT   oxidoreductase, for bioconversion of steroid substances to their 11 alpha
PT   hydroxy counterparts in heterologous cells.
XX
PS   Example 10; Page 57; 181pp; English.
XX
CC   The present invention relates to novel cytochrome P450-like enzyme
CC   (Aspergillus ochraceus 11 alpha hydroxylase protein), oxidoreductases and
CC   polynucleotides encoding such proteins. Host cells comprising the
CC   sequences of the invention are useful for making one or more enzymes from
CC   the metabolic pathway for the synthesis of sitosterol to eplerenone. They
CC   are useful for selective oxidation of a compound to an hydroxylated
CC   product. Compositions of the invention are useful for producing spores
CC   from A. ochraceus, A. niger, A. nidulans, Rhizopus oryzae, R. stolonifer,
CC   R. arrhizus Trichothecium roseum, Fusarium oxysporum and M. olivaceum
CC   etc, preferably to produce spores from A. ochraceus. Sequences of the
CC   invention are useful in bioconversion of steroid substances to their 11
CC   alpha-hydroxy counterparts. They are also used in cell therapy. The
CC   present sequence is A. ochraceus cytochrome p450 conserved haem binding
CC   motif
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 33;  DB 5;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              ||||||||||
Db          1 FXXGXXXCXG 10


RESULT 10
ABU08352
ID   ABU08352 standard; peptide; 10 AA.
XX
AC   ABU08352;
XX
DT   30-MAY-2003  (first entry)
XX
DE   Conserved haem-binding motif found in eukaryotic P450 monooxygenases.
XX
KW   Cytochrome P450; (+)-delta-cadinene 8-hydroxylase; CYP706B1;
KW   biosynthesis of gossypol; sesquiterpene; cotton seed; cotton cultivate;
KW   sesquiterpenoid; livestock feed; monooxygenase; haem-binding motif.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 2. .3
FT                   /note= "Any amino acid"
FT   Misc-difference 5
FT                   /note= "Any amino acid"
FT   Misc-difference 7
FT                   /note= "Any amino acid"
FT   Misc-difference 9
FT                   /note= "Any amino acid"
XX
PN   US2002187538-A1.
XX
PD   12-DEC-2002.
XX
PF   07-FEB-2002; 2002US-00067534.
XX
PR   07-FEB-2001; 2001US-0267160P.
XX
PA   (ESSE/) ESSENBERG M K.
PA   (CHEN/) CHEN X.
PA   (LUOP/) LUO P.
PA   (WANG/) WANG Y.
XX
PI   Essenberg MK,  Chen X,  Luo P,  Wang Y;
XX
DR   WPI; 2003-341036/32.
XX
PT   Novel cotton (+)-gamma-cadinene 8-hydroxylase polypeptide designated as
PT   CYP706B1, useful as target for suppression of biosynthesis of gossypol
PT   formation in cotton seeds.
XX
PS   Example 1; Page 6; 26pp; English.
XX
CC   The present invention relates to the isolation of cotton (+)-delta-
CC   cadinene 8-hydroxylase (designated as CYP706B1), and the polynucleotide
CC   sequence encoding it. The CYP706B1 protein is a cytochrome P450 which is
CC   useful as a target for suppression of the biosynthesis of gossypol and
CC   related sesquiterpenes in cotton seeds through genetic engineering
```

```
CC   techniques. The polynucleotide sequence encoding CYP706B1 is useful in
CC   suppression of the biosynthesis of gossypol and related sesquiterpenes in
CC   cotton seeds, where the polynucleotide sequence is expressed in antisense
CC   or sense orientation as a perfect match to the native gene whose
CC   expression is sought to be suppressed. The polynucleotide sequence of the
CC   invention is useful for producing cotton cultivates which avoid the
CC   presence of sesquiterpenoids in their seeds, and for producing cotton
CC   seed product which is suitable for use as a feed for both livestock and
CC   humans. The present sequence represents a highly conserved haem-binding
CC   motif found in eukaryotic P450 monooxygenases
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 33;  DB 6;  Length 10;
  Best Local Similarity   90.0%;  Pred. No. 3.5e+02;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |||||:||||
Db          1 FXXGXRXCXG 10


RESULT 11
ABR42010
ID   ABR42010 standard; peptide; 10 AA.
XX
AC   ABR42010;
XX
DT   11-AUG-2003  (first entry)
XX
DE   Cytochrome P450 conserved motif.
XX
KW   Cytochrome P450; EthB; ether; ethyl tert-butyl ether; ETBE; degradation;
KW   bioremediation; soil decontamination; pollutant; biosensor;
KW   Rhodococcus ruber.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 3
FT                   /note= "any amino acid"
FT   Misc-difference 5
FT                   /note= "any amino acid"
FT   Misc-difference 7
FT                   /note= "any amino acid"
FT   Misc-difference 9
FT                   /note= "any amino acid"
XX
PN   EP1270722-A1.
XX
PD   02-JAN-2003.
XX
PF   22-JUN-2001; 2001EP-00401667.
XX
PR   22-JUN-2001; 2001EP-00401667.
XX
PA   (INSP ) INST PASTEUR.
PA   (INSF ) INST FRANCAIS DU PETROLE.
XX
PI   Chauvaux S,  Miras I,  Beguin P;
XX
DR   WPI; 2003-334761/32.
XX
PT   New Rhodococcus ruber nucleic acid, useful for transforming bacteria for
PT   depolluting soil contaminated with ethyl tert-butyl ether comprises the
PT   cytochrome P-450 gene cluster involved in cleavage of ethyl tert-butyl
PT   ether.
XX
PS   Example 5; Page 10; 58pp; English.
XX
CC   The present sequence is a consensus motif of chytochrome P450s, and
CC   includes a Cys residue that is strictly conserved in all cytochrome
CC   P450s. The motif has been identified in the EthB protein (see ABR42001)
CC   of Rhodococcus ruber strain CNCM I-1889 cells. EthB is an ethyl tert-
CC   butyl ether (ETBE)-induced cytochrome P450. It forms part of a cytochrome
CC   P450 system involved in the cleavage of ether fuel additives. Recombinant
CC   host cells comprising a vector encoding eth genes, and which are capable
CC   of ETBE degradation or mineralization, can be used for bioremediation of
CC   a contaminated soil or effluent
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 33;  DB 6;  Length 10;
  Best Local Similarity   80.0%;  Pred. No. 3.5e+02;
  Matches    8;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |:|||:||||
Db          1 FGXGXHXCXG 10


RESULT 12
AAE35263
ID   AAE35263 standard; peptide; 10 AA.
XX
AC   AAE35263;
XX
DT   28-MAY-2003  (first entry)
XX
DE   Cytochrome P450 putative family signature peptide #2.
XX
KW   Human; retinoic acid; vitamin A; cytochrome P450; actinic keratosis;
KW   P450RAI; cancer; psoriasis; acne; ichthyosis; gene therapy; vaccine;
KW   enzyme.
XX
OS   Unidentified.
XX
PN   WO200295034-A2.
XX
PD   28-NOV-2002.
XX
PF   23-MAY-2002; 2002WO-CA000758.
XX
PR   23-MAY-2001; 2001US-0292531P.
XX
PA   (CYTO-) CYTOCHROMA INC.
XX
PI   Wisniewski J,  Petkovich PM,  Ramshaw H;
XX
DR   WPI; 2003-112152/10.
XX
PT   New isolated P450RAI-3 nucleic acid molecule and polypeptide, useful for
PT   diagnosing, preventing, or treating disorders with aberrant expression or
PT   activity of the P450RAI-3, such as cancer, actinic keratosis, acne and
PT   psoriasis.
XX
PS   Example 1; Page 72; 231pp; English.
XX
CC   The invention relates to retinoic acid metabolising cytochrome P450,
CC   P450RAI polypeptides and polynucleotides. Methods and compositions of the
CC   invention are useful for diagnosing, preventing, ameliorating and/or
CC   treating disorders associated with the aberrant expression or activity of
CC   the P450RAI such as diseases related to vitamin A and retinoic acid
CC   metabolism, e.g. cancer, actinic keratosis, psoriasis, acne and
CC   ichthyosis. The invention is useful in gene therapy and as vaccine. The
CC   present sequence is cytochrome P450 putative family signature peptide.
CC   This peptide is used in the exemplification of the invention
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 33;  DB 6;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 3.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
```

```
Db          1 FGGGARSCLG 10


RESULT 13
AAE35262
ID   AAE35262 standard; peptide; 10 AA.
XX
AC   AAE35262;
XX
DT   28-MAY-2003  (first entry)
XX
DE   Cytochrome P450 putative family signature peptide #1.
XX
KW   Human; retinoic acid; vitamin A; cytochrome P450; actinic keratosis;
KW   P450RAI; cancer; psoriasis; acne; ichthyosis; gene therapy; vaccine;
KW   enzyme.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 2. .3
FT                   /label= Unknown
FT   Misc-difference 5. .7
FT                   /label= Unknown
FT   Misc-difference 9
FT                   /label= Unknown
XX
PN   WO200295034-A2.
XX
PD   28-NOV-2002.
XX
PF   23-MAY-2002; 2002WO-CA000758.
XX
PR   23-MAY-2001; 2001US-0292531P.
XX
PA   (CYTO-) CYTOCHROMA INC.
XX
PI   Wisniewski J,  Petkovich PM,  Ramshaw H;
XX
DR   WPI; 2003-112152/10.
XX
PT   New isolated P450RAI-3 nucleic acid molecule and polypeptide, useful for
PT   diagnosing, preventing, or treating disorders with aberrant expression or
PT   activity of the P450RAI-3, such as cancer, actinic keratosis, acne and
PT   psoriasis.
XX
PS   Example 1; Page 72; 231pp; English.
XX
CC   The invention relates to retinoic acid metabolising cytochrome P450,
CC   P450RAI polypeptides and polynucleotides. Methods and compositions of the
CC   invention are useful for diagnosing, preventing, ameliorating and/or
CC   treating disorders associated with the aberrant expression or activity of
CC   the P450RAI such as diseases related to vitamin A and retinoic acid
CC   metabolism, e.g. cancer, actinic keratosis, psoriasis, acne and
CC   ichthyosis. The invention is useful in gene therapy and as vaccine. The
CC   present sequence is cytochrome P450 putative family signature peptide.
CC   This peptide is used in the exemplification of the invention
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 33;  DB 6;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              ||||||||||
Db          1 FXXGXXXCXG 10


RESULT 14
ADC56716
ID   ADC56716 standard; peptide; 10 AA.
XX
AC   ADC56716;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Peptide 1 related to human vitamin-D3 hydroxylase.
XX
KW   vitamin-D3 hydroxylase; VDDH; antipsoriatic; osteopathic; hypercalcaemia;
KW   gene therapy; antisense; hypoparathyroidism; rickets; psoriasis;
KW   osteoporosis; renal insufficiency.
XX
OS   Unidentified.
XX
PN   JP2003047479-A.
XX
PD   18-FEB-2003.
XX
PF   08-AUG-2001; 2001JP-00241396.
XX
PR   08-AUG-2001; 2001JP-00241396.
XX
PA   (MOCH ) MOCHIDA PHARM CO LTD.
PA   (HERI-) HERIKKUSU KENKYUSHO KK.
XX
DR   WPI; 2003-508705/48.
XX
PT   Novel vitamin-D3 hydroxylase protein, useful for identifying modulators
PT   of the vitamin-D3 hydroxylation activity of the enzyme.
XX
PS   Example 1; Page 10; 18pp; Japanese.
XX
CC   The invention relates to a novel vitamin-D3 hydroxylase (VDDH) protein.
CC   The polypeptide of the invention demonstrates antipsoriatic and
CC   osteopathic activities and may be useful for identifying compounds which
CC   modulate the activity or inhibit the expression of vitamin-D3 hydroxylase
CC   and thus for treating or preventing hypercalcaemia. Furthermore, the
CC   polypeptide may be utilised during gene and antisense therapy, as well as
CC   for treating activated vitamin-D3 deficiency disorders such as
CC   hypoparathyroidism, rickets, psoriasis, osteoporosis and renal
CC   insufficiency. The current sequence is that of the peptide 1 of the
CC   invention which is related to human vitamin-D3 hydroxylase.
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 33;  DB 7;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 3.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGHGVRSCIG 10


RESULT 15
ADC98703
ID   ADC98703 standard; peptide; 10 AA.
XX
AC   ADC98703;
XX
DT   01-JAN-2004  (first entry)
XX
DE   Cytochrome P450 haem ligand signature motif peptide.
XX
KW   ecdysteroid synthesis inhibitor; insecticidal; cytochrome P450;
KW   haem ligand motif.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 2. .3
FT                   /label= OTHER
FT                   /note= "OTHER = Any amino acid"
FT   Misc-difference 5. .7
FT                   /label= OTHER
```

```
FT                   /note= "OTHER = Any amino acid"
FT   Misc-difference 9
FT                   /label= OTHER
FT                   /note= "OTHER = Any amino acid"
XX
PN   US2003100025-A1.
XX
PD   29-MAY-2003.
XX
PF   06-SEP-2002; 2002US-00236433.
XX
PR   07-SEP-2001; 2001US-0317890P.
PR   07-SEP-2001; 2001US-0318006P.
XX
PA   (OCON/) O'CONNOR M.
PA   (GILB/) GILBERT L I.
PA   (WARR/) WARREN J T.
XX
PI   O'connor M,  Gilbert LI,  Warren JT;
XX
DR   WPI; 2003-777312/73.
XX
PT   New polypeptide useful for identifying inhibitors of ecdysteroid
PT   synthesis, determining insecticidal properties, and making an
PT   insecticide.
XX
PS   Disclosure; Page 3; 21pp; English.
XX
CC   The invention relates to a novel method for identifying inhibitors of
CC   ecdysteroid synthesis comprising contacting an ecdysteroid biosynthetic
CC   enzyme with a candidate inhibitor molecule and determining whether or not
CC   said inhibitor inhibits enzymatic activity. The enzyme comprises a novel
CC   isolated polypeptide of CYP302a1, CYP306a1, CYP307a1, CYP314a1 or
CC   CYP315a1. The method of the invention may be useful for determining
CC   whether or not a molecule has insecticidal properties and for making
CC   inhibitors of ecdysteroid synthesis. The current sequence is that of the
CC   cytochrome P450 haem ligand signature motif peptide of the invention.
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 33;  DB 7;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              ||||||||||
Db          1 FXXGXXXCXG 10


RESULT 16
AEB16917
ID   AEB16917 standard; peptide; 10 AA.
XX
AC   AEB16917;
XX
DT   08-SEP-2005  (first entry)
XX
DE   LUT1 cytochrome P450 monooxygenase cysteine motif peptide, SEQ ID NO: 15.
XX
KW   Pigment; metabolic engineering; antioxidant; transgenic plant;
KW   cytochrome P450.
XX
OS   Arabidopsis thaliana.
XX
PN   US2005150002-A1.
XX
PD   07-JUL-2005.
XX
PF   02-JAN-2004; 2004US-00751235.
XX
PR   02-JAN-2004; 2004US-00751235.
XX
PA   (DELL/) DELLAPENNA D.
PA   (TIAN/) TIAN L.
PA   (KIMJ/) KIM J.
XX
PI   Dellapenna D,  Tian L,  Kim J;
XX
DR   WPI; 2005-487984/49.
XX
PT   New expression vector comprising a nucleic acid sequence encoding a
PT   polypeptide having monooxygenase P450 activity, useful in altering the
PT   carotenoid production in a plant for enhancing production of specific
PT   carotenoid compounds.
XX
PS   Disclosure; SEQ ID NO 15; 135pp; English.
XX
CC   The present invention relates to genes, proteins and methods comprising
CC   carotenoid monooxygenases in the cytochrome P450 family. The invention
CC   also relates to altering carotenoid ratios in plants and microorganisms
CC   using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC   invention is useful in altering the carotenoid production in a plant for
CC   enhancing production of specific carotenoid compounds that are potent
CC   antioxidants. The present sequence is thale cress LUT1 cytochrome P450
CC   monooxygenase cysteine motif peptide.
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 33;  DB 9;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 3.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FSGGPRKCVG 10


RESULT 17
AEB16916
ID   AEB16916 standard; peptide; 10 AA.
XX
AC   AEB16916;
XX
DT   08-SEP-2005  (first entry)
XX
DE   Cytochrome P450 monooxygenase cysteine consensus motif, SEQ ID NO: 14.
XX
KW   Pigment; metabolic engineering; antioxidant; transgenic plant;
KW   cytochrome P450.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 2. .3
FT                   /note= "Xaa may be any naturally occurring amino acid"
FT   Misc-difference 5. .7
FT                   /note= "Xaa may be any naturally occurring amino acid"
FT   Misc-difference 9
FT                   /note= "Xaa may be any naturally occurring amino acid"
XX
PN   US2005150002-A1.
XX
PD   07-JUL-2005.
XX
PF   02-JAN-2004; 2004US-00751235.
XX
PR   02-JAN-2004; 2004US-00751235.
XX
PA   (DELL/) DELLAPENNA D.
PA   (TIAN/) TIAN L.
PA   (KIMJ/) KIM J.
XX
PI   Dellapenna D,  Tian L,  Kim J;
XX
DR   WPI; 2005-487984/49.
XX
```

PT New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.
XX
PS Claim 7; SEQ ID NO 14; 135pp; English.
XX
CC The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is a cytochrome P450 monooxygenase
CC cysteine consensus motif peptide.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 33; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
||||||||||
Db 1 FXXGXXXCXG 10

RESULT 18
AAR30867
ID AAR30867 standard; protein; 11 AA.
XX
AC AAR30867;
XX
DT 25-MAR-2003 (revised)
DT 12-MAY-1993 (first entry)
XX
DE Consensus haem-binding domain of cyt-P450 in plants.
XX
KW Dihydrokaempferol; flavonoid; pigmentation; colour; amplification;
KW cytochrome P450; advocado; petunia; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 4
FT /label= Ala, Ser
FT Misc-difference 6
FT /label= Lys, Arg
FT Misc-difference 8
FT /label= Ile, Gly
XX
PN EP522880-A2.
XX
PD 13-JAN-1993.
XX
PF 10-JUL-1992; 92EP-00306379.
XX
PR 11-JUL-1991; 91AU-00007173.
PR 17-FEB-1992; 92AU-00000923.
XX
PA (ITFL-) INT FLOWER DEV PTY LTD.
XX
PI Holton TA, Cornish EC, Kovacic F, Tanaka Y, Lester DR;
XX
DR WPI; 1993-010688/02.
XX
PT Nucleic acid sequence encoding a di:hydro:kaempferol-hydroxylating enzyme
PT - e.g. cytochrome P450 introduced into transgenic plants for controlling
PT flavonoid pigmentation in plants and organisms.
XX
PS Disclosure; Page 13; 66pp; English.
XX
CC The consensus sequence of haem-binding domains for advocado and petunia
CC cytochrome P450 homologues pCGP142 and pCGP147 was used to design
CC synthetic PCR primers. The primers may be used in PCR for amplification
CC of petal cytochrome P450 homologues. See also AAR30868-75. (Updated on 25
CC -MAR-2003 to correct PN field.)
XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 33; DB 2; Length 11;
Best Local Similarity 70.0%; Pred. No. 3.9e+02;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|:|||:||:|
Db 2 FGXGXRXCPG 11

RESULT 19
AAR40877
ID AAR40877 standard; protein; 11 AA.
XX
AC AAR40877;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 28-MAR-1994 (first entry)
XX
DE SSP for flavonoid-3',5'-hydroxylase gene product.
XX
KW Flavonoid-3',5'-hydoxylase; transformation; plants; petunia; rose;
KW tobacco; pigment alteration; blue; SSP; single specific primer; PCR;
KW polymerase chain reaction; amplification; expression; ss.
XX
OS Petunia x hybrida.
XX
PN WO9318155-A1.
XX
PD 16-SEP-1993.
XX
PF 20-NOV-1992; 92WO-JP001520.
XX
PR 02-MAR-1992; 92JP-00044963.
XX
PA (KYOW) KYOWA HAKKO KOGYO CO LTD.
XX
PI Kikuchi Y, Kiyokawa S, Shimada Y, Ohbayashi M, Shimada R;
PI Okinaka Y;
XX
DR WPI; 1993-303469/38.
DR N-PSDB; AAQ47878.
XX
PT Gene coding for flavonoid-3',5'-hydroxylase of petunia petals - used to
PT transform plants e.g. petunia, rose or tobacco to give bluer flower
PT colour and altered pigment pattern.
XX
PS Claim 11; Page 58; 82pp; Japanese.
XX
CC Insertion of the sequences (AAQ47840-42) into plants such as rose,
CC petunia, tobacco and carnation, using a suitable vector such as
CC agrobacterium, give transformed plants which express the gene, resulting
CC in petals with a bluer colour than normal, and/or pigmentation patterns
CC which do not occur naturally. The sequences were amplified using primers
CC (AAQ47843-70). Related single specific primers using a gene sequence
CC coding for the haem-binding region of cytochrome P450 are shown in
CC (AAQ47871-Q47903). (Updated on 25-MAR-2003 to correct PN field.) (Updated
CC on 24-OCT-2003 to standardise OS field)
XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 33; DB 2; Length 11;
Best Local Similarity 40.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10

```
                |::|:::|:|
Db          2 FGGGPRRCPG 11


RESULT 20
AAR40874
ID   AAR40874 standard; protein; 11 AA.
XX
AC   AAR40874;
XX
DT   24-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   28-MAR-1994  (first entry)
XX
DE   SSP for flavonoid-3',5'-hydroxylase gene.
XX
KW   Flavonoid-3',5'-hydoxylase; transformation; plants; petunia; rose;
KW   tobacco; pigment alteration; blue; SSP; single specific primer; PCR;
KW   polymerase chain reaction; amplification; expression; ss.
XX
OS   Petunia x hybrida.
XX
PN   WO9318155-A1.
XX
PD   16-SEP-1993.
XX
PF   20-NOV-1992;   92WO-JP001520.
XX
PR   02-MAR-1992;   92JP-00044963.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO CO LTD.
XX
PI   Kikuchi Y,  Kiyokawa S,  Shimada Y,  Ohbayashi M,  Shimada R;
PI   Okinaka Y;
XX
DR   WPI; 1993-303469/38.
DR   N-PSDB; AAQ47875.
XX
PT   Gene coding for flavonoid-3',5'-hydroxylase of petunia petals - used to
PT   transform plants e.g. petunia, rose or tobacco to give bluer flower
PT   colour and altered pigment pattern.
XX
PS   Claim 11; Page 57; 82pp; Japanese.
XX
CC   Insertion of the sequences (AAQ47840-42) into plants such as rose,
CC   petunia, tobacco and carnation, using a suitable vector such as
CC   agrobacterium, give transformed plants which express the gene, resulting
CC   in petals with a bluer colour than normal, and/or pigmentation patterns
CC   which do not occur naturally. The sequences were amplified using primers
CC   (AAQ47843-70). Related single specific primers using a gene sequence
CC   coding for the haem-binding region of cytochrome P450 are shown in
CC   (AAQ47871-Q47903). (Updated on 25-MAR-2003 to correct PN field.) (Updated
CC   on 24-OCT-2003 to standardise OS field)
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 33;  DB 2;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGSGRRSCPG 11


RESULT 21
AAR40876
ID   AAR40876 standard; protein; 11 AA.
XX
AC   AAR40876;
XX
DT   24-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   28-MAR-1994  (first entry)
XX
DE   SSP for flavonoid-3',5'-hydroxylase gene.
XX
KW   Flavonoid-3',5'-hydoxylase; transformation; plants; petunia; rose;
KW   tobacco; pigment alteration; blue; SSP; single specific primer; PCR;
KW   polymerase chain reaction; amplification; expression; ss.
XX
OS   Petunia x hybrida.
XX
PN   WO9318155-A1.
XX
PD   16-SEP-1993.
XX
PF   20-NOV-1992;   92WO-JP001520.
XX
PR   02-MAR-1992;   92JP-00044963.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO CO LTD.
XX
PI   Kikuchi Y,  Kiyokawa S,  Shimada Y,  Ohbayashi M,  Shimada R;
PI   Okinaka Y;
XX
DR   WPI; 1993-303469/38.
DR   P-PSDB; AAR40876.
XX
PT   Gene coding for flavonoid-3',5'-hydroxylase of petunia petals - used to
PT   transform plants e.g. petunia, rose or tobacco to give bluer flower
PT   colour and altered pigment pattern.
XX
PS   Claim 9; Page 57; 82pp; Japanese.
XX
CC   Insertion of the sequences (AAQ47840-42) into plants such as rose,
CC   petunia, tobacco and carnation, using a suitable vector such as
CC   agrobacterium, give transformed plants which express the gene, resulting
CC   in petals with a bluer colour than normal, and/or pigmentation patterns
CC   which do not occur naturally. The sequences were amplified using primers
CC   (AAQ47843-70). Related single specific primers using a gene sequence
CC   coding for the haem-binding region of cytochrome P450 are shown in
CC   (AAQ47871-Q47903). (Updated on 25-MAR-2003 to correct PN field.) (Updated
CC   on 24-OCT-2003 to standardise OS field)
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 33;  DB 2;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGVGLRMCPG 11


RESULT 22
AAR41633
ID   AAR41633 standard; protein; 11 AA.
XX
AC   AAR41633;
XX
DT   24-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   28-MAR-1994  (first entry)
XX
DE   SSP for flavonoid-3',5'-hydroxylase gene product.
XX
KW   Flavonoid-3',5'-hydoxylase; transformation; plants; petunia; rose;
KW   tobacco; pigment alteration; blue; SSP; single specific primer; PCR;
KW   polymerase chain reaction; amplification; expression.
XX
OS   Petunia x hybrida.
XX
PN   WO9318155-A1.
XX
```

```
PD   16-SEP-1993.
XX
PF   20-NOV-1992;   92WO-JP001520.
XX
PR   02-MAR-1992;   92JP-00044963.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO CO LTD.
XX
PI   Kikuchi Y,  Kiyokawa S,  Shimada Y,  Ohbayashi M,  Shimada R;
PI   Okinaka Y;
XX
DR   WPI; 1993-303469/38.
DR   N-PSDB; AAQ47883.
XX
PT   Gene coding for flavonoid-3',5'-hydroxylase of petunia petals - used to
PT   transform plants e.g. petunia, rose or tobacco to give bluer flower
PT   colour and altered pigment pattern.
XX
PS   Claim 11; Page 60; 82pp; Japanese.
XX
CC   Insertion of the sequences (AAQ47840-42) into plants such as rose,
CC   petunia, tobacco and carnation, using a suitable vector such as
CC   agrobacterium, give transformed plants which express the gene, resulting
CC   in petals with a bluer colour than normal, and/or pigmentation patterns
CC   which do not occur naturally. The sequences were amplified using primers
CC   (AAQ47843-70). Related single specific primers using a gene sequence
CC   coding for the haem-binding region of cytochrome P450 are shown in
CC   (AAQ47871-Q47903). (Updated on 25-MAR-2003 to correct PN field.) (Updated
CC   on 24-OCT-2003 to standardise OS field)
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 33;  DB 2;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGSGFCSCPG 11


RESULT 23
AAR40872
ID   AAR40872 standard; protein; 11 AA.
XX
AC   AAR40872;
XX
DT   24-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   28-MAR-1994  (first entry)
XX
DE   SSP for flavonoid-3',5'-hydroxylase gene product.
XX
KW   Flavonoid-3',5'-hydoxylase; transformation; plants; petunia; rose;
KW   tobacco; pigment alteration; blue; SSP; single specific primer; PCR;
KW   polymerase chain reaction; amplification; expression.
XX
OS   Petunia x hybrida.
XX
PN   WO9318155-A1.
XX
PD   16-SEP-1993.
XX
PF   20-NOV-1992;   92WO-JP001520.
XX
PR   02-MAR-1992;   92JP-00044963.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO CO LTD.
XX
PI   Kikuchi Y,  Kiyokawa S,  Shimada Y,  Ohbayashi M,  Shimada R;
PI   Okinaka Y;
XX
DR   WPI; 1993-303469/38.
DR   N-PSDB; AAQ47873.
XX
PT   Gene coding for flavonoid-3',5'-hydroxylase of petunia petals - used to
PT   transform plants e.g. petunia, rose or tobacco to give bluer flower
PT   colour and altered pigment pattern.
XX
PS   Claim 11; Page 56; 82pp; Japanese.
XX
CC   Insertion of the sequences (AAQ47840-42) into plants such as rose,
CC   petunia, tobacco and carnation, using a suitable vector such as
CC   agrobacterium, give transformed plants which express the gene, resulting
CC   in petals with a bluer colour than normal, and/or pigmentation patterns
CC   which do not occur naturally. The sequences were amplified using primers
CC   (AAQ47843-70). Related single specific primers using a gene sequence
CC   coding for the haem-binding region of cytochrome P450 are shown in
CC   (AAQ47871-Q47903). (Updated on 25-MAR-2003 to correct PN field.) (Updated
CC   on 24-OCT-2003 to standardise OS field)
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 33;  DB 2;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGAGRRICPG 11


RESULT 24
AAR40879
ID   AAR40879 standard; protein; 11 AA.
XX
AC   AAR40879;
XX
DT   24-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   28-MAR-1994  (first entry)
XX
DE   SSP for flavonoid-3',5'-hydroxylase gene product.
XX
KW   Flavonoid-3',5'-hydoxylase; transformation; plants; petunia; rose;
KW   tobacco; pigment alteration; blue; SSP; single specific primer; PCR;
KW   polymerase chain reaction; amplification; expression.
XX
OS   Petunia x hybrida.
XX
PN   WO9318155-A1.
XX
PD   16-SEP-1993.
XX
PF   20-NOV-1992;   92WO-JP001520.
XX
PR   02-MAR-1992;   92JP-00044963.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO CO LTD.
XX
PI   Kikuchi Y,  Kiyokawa S,  Shimada Y,  Ohbayashi M,  Shimada R;
PI   Okinaka Y;
XX
DR   WPI; 1993-303469/38.
DR   N-PSDB; AAQ47880.
XX
PT   Gene coding for flavonoid-3',5'-hydroxylase of petunia petals - used to
PT   transform plants e.g. petunia, rose or tobacco to give bluer flower
PT   colour and altered pigment pattern.
XX
PS   Claim 11; Page 59; 82pp; Japanese.
XX
CC   Insertion of the sequences (AAQ47840-42) into plants such as rose,
CC   petunia, tobacco and carnation, using a suitable vector such as
CC   agrobacterium, give transformed plants which express the gene, resulting
CC   in petals with a bluer colour than normal, and/or pigmentation patterns
```

```
CC   which do not occur naturally. The sequences were amplified using primers
CC   (AAQ47843-70). Related single specific primers using a gene sequence
CC   coding for the haem-binding region of cytochrome P450 are shown in
CC   (AAQ47871-Q47903). (Updated on 25-MAR-2003 to correct PN field.) (Updated
CC   on 24-OCT-2003 to standardise OS field)
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 33;  DB 2;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGVGPKMCPG 11


RESULT 25
AAR41632
ID   AAR41632 standard; protein; 11 AA.
XX
AC   AAR41632;
XX
DT   24-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   28-MAR-1994  (first entry)
XX
DE   SSP for flavonoid-3',5'-hydroxylase gene product.
XX
KW   Flavonoid-3',5'-hydoxylase; transformation; plants; petunia; rose;
KW   tobacco; pigment alteration; blue; SSP; single specific primer; PCR;
KW   polymerase chain reaction; amplification; expression.
XX
OS   Petunia x hybrida.
XX
PN   WO9318155-A1.
XX
PD   16-SEP-1993.
XX
PF   20-NOV-1992;   92WO-JP001520.
XX
PR   02-MAR-1992;   92JP-00044963.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO CO LTD.
XX
PI   Kikuchi Y,  Kiyokawa S,  Shimada Y,  Ohbayashi M,  Shimada R;
PI   Okinaka Y;
XX
DR   WPI; 1993-303469/38.
DR   N-PSDB; AAQ47882.
XX
PT   Gene coding for flavonoid-3',5'-hydroxylase of petunia petals - used to
PT   transform plants e.g. petunia, rose or tobacco to give bluer flower
PT   colour and altered pigment pattern.
XX
PS   Claim 11; Page 60; 82pp; Japanese.
XX
CC   Insertion of the sequences (AAQ47840-42) into plants such as rose,
CC   petunia, tobacco and carnation, using a suitable vector such as
CC   agrobacterium, give transformed plants which express the gene, resulting
CC   in petals with a bluer colour than normal, and/or pigmentation patterns
CC   which do not occur naturally. The sequences were amplified using primers
CC   (AAQ47843-70). Related single specific primers using a gene sequence
CC   coding for the haem-binding region of cytochrome P450 are shown in
CC   (AAQ47871-Q47903). (Updated on 25-MAR-2003 to correct PN field.) (Updated
CC   on 24-OCT-2003 to standardise OS field)
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 33;  DB 2;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGFGPRKCVG 11


RESULT 26
AAR41634
ID   AAR41634 standard; protein; 11 AA.
XX
AC   AAR41634;
XX
DT   24-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   28-MAR-1994  (first entry)
XX
DE   SSP for flavonoid-3',5'-hydroxylase gene product.
XX
KW   Flavonoid-3',5'-hydoxylase; transformation; plants; petunia; rose;
KW   tobacco; pigment alteration; blue; SSP; single specific primer; PCR;
KW   polymerase chain reaction; amplification; expression.
XX
OS   Petunia x hybrida.
XX
PN   WO9318155-A1.
XX
PD   16-SEP-1993.
XX
PF   20-NOV-1992;   92WO-JP001520.
XX
PR   02-MAR-1992;   92JP-00044963.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO CO LTD.
XX
PI   Kikuchi Y,  Kiyokawa S,  Shimada Y,  Ohbayashi M,  Shimada R;
PI   Okinaka Y;
XX
DR   WPI; 1993-303469/38.
DR   N-PSDB; AAQ47883.
XX
PT   Gene coding for flavonoid-3',5'-hydroxylase of petunia petals - used to
PT   transform plants e.g. petunia, rose or tobacco to give bluer flower
PT   colour and altered pigment pattern.
XX
PS   Claim 11; Page 61; 82pp; Japanese.
XX
CC   Insertion of the sequences (AAQ47840-42) into plants such as rose,
CC   petunia, tobacco and carnation, using a suitable vector such as
CC   agrobacterium, give transformed plants which express the gene, resulting
CC   in petals with a bluer colour than normal, and/or pigmentation patterns
CC   which do not occur naturally. The sequences were amplified using primers
CC   (AAQ47843-70). Related single specific primers using a gene sequence
CC   coding for the haem-binding region of cytochrome P450 are shown in
CC   (AAQ47871-Q47903). (Updated on 25-MAR-2003 to correct PN field.) (Updated
CC   on 24-OCT-2003 to standardise OS field)
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 33;  DB 2;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGAGRRICAG 11


RESULT 27
AAR40873
ID   AAR40873 standard; protein; 11 AA.
XX
AC   AAR40873;
XX
DT   24-OCT-2003  (revised)
```

```
DT   25-MAR-2003  (revised)
DT   28-MAR-1994  (first entry)
XX
DE   SSP for flavonoid-3',5'-hydroxylase gene product.
XX
KW   Flavonoid-3',5'-hydoxylase; transformation; plants; petunia; rose;
KW   tobacco; pigment alteration; blue; SSP; single specific primer; PCR;
KW   polymerase chain reaction; amplification; expression.
XX
OS   Petunia x hybrida.
XX
PN   WO9318155-A1.
XX
PD   16-SEP-1993.
XX
PF   20-NOV-1992;   92WO-JP001520.
XX
PR   02-MAR-1992;   92JP-00044963.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO CO LTD.
XX
PI   Kikuchi Y,  Kiyokawa S,  Shimada Y,  Ohbayashi M,  Shimada R;
PI   Okinaka Y;
XX
DR   WPI; 1993-303469/38.
DR   N-PSDB; AAQ47874.
XX
PT   Gene coding for flavonoid-3',5'-hydroxylase of petunia petals - used to
PT   transform plants e.g. petunia, rose or tobacco to give bluer flower
PT   colour and altered pigment pattern.
XX
PS   Claim 11; Page 56; 82pp; Japanese.
XX
CC   Insertion of the sequences (AAQ47840-42) into plants such as rose,
CC   petunia, tobacco and carnation, using a suitable vector such as
CC   agrobacterium, give transformed plants which express the gene, resulting
CC   in petals with a bluer colour than normal, and/or pigmentation patterns
CC   which do not occur naturally. The sequences were amplified using primers
CC   (AAQ47843-70). Related single specific primers using a gene sequence
CC   coding for the haem-binding region of cytochrome P450 are shown in
CC   (AAQ47871-Q47903). (Updated on 25-MAR-2003 to correct PN field.) (Updated
CC   on 24-OCT-2003 to standardise OS field)
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 33;  DB 2;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGTGRRICPG 11


RESULT 28
AAR40878
ID   AAR40878 standard; protein; 11 AA.
XX
AC   AAR40878;
XX
DT   24-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   28-MAR-1994  (first entry)
XX
DE   SSP for flavonoid-3',5'-hydroxylase gene product.
XX
KW   Flavonoid-3',5'-hydoxylase; transformation; plants; petunia; rose;
KW   tobacco; pigment alteration; blue; SSP; single specific primer; PCR;
KW   polymerase chain reaction; amplification; expression.
XX
OS   Petunia x hybrida.
XX
PN   WO9318155-A1.
XX
PD   16-SEP-1993.
XX
PF   20-NOV-1992;   92WO-JP001520.
XX
PR   02-MAR-1992;   92JP-00044963.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO CO LTD.
XX
PI   Kikuchi Y,  Kiyokawa S,  Shimada Y,  Ohbayashi M,  Shimada R;
PI   Okinaka Y;
XX
DR   WPI; 1993-303469/38.
DR   N-PSDB; AAQ47879.
XX
PT   Gene coding for flavonoid-3',5'-hydroxylase of petunia petals - used to
PT   transform plants e.g. petunia, rose or tobacco to give bluer flower
PT   colour and altered pigment pattern.
XX
PS   Claim 11; Page 58; 82pp; Japanese.
XX
CC   Insertion of the sequences (AAQ47840-42) into plants such as rose,
CC   petunia, tobacco and carnation, using a suitable vector such as
CC   agrobacterium, give transformed plants which express the gene, resulting
CC   in petals with a bluer colour than normal, and/or pigmentation patterns
CC   which do not occur naturally. The sequences were amplified using primers
CC   (AAQ47843-70). Related single specific primers using a gene sequence
CC   coding for the haem-binding region of cytochrome P450 are shown in
CC   (AAQ47871-Q47903). (Updated on 25-MAR-2003 to correct PN field.) (Updated
CC   on 24-OCT-2003 to standardise OS field)
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 33;  DB 2;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGVGRRSCPG 11


RESULT 29
AAR40880
ID   AAR40880 standard; protein; 11 AA.
XX
AC   AAR40880;
XX
DT   24-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   28-MAR-1994  (first entry)
XX
DE   SSP for flavonoid-3',5'-hydroxylase gene product.
XX
KW   Flavonoid-3',5'-hydoxylase; transformation; plants; petunia; rose;
KW   tobacco; pigment alteration; blue; SSP; single specific primer; PCR;
KW   polymerase chain reaction; amplification; expression.
XX
OS   Petunia x hybrida.
XX
PN   WO9318155-A1.
XX
PD   16-SEP-1993.
XX
PF   20-NOV-1992;   92WO-JP001520.
XX
PR   02-MAR-1992;   92JP-00044963.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO CO LTD.
XX
PI   Kikuchi Y,  Kiyokawa S,  Shimada Y,  Ohbayashi M,  Shimada R;
PI   Okinaka Y;
XX
```

```
DR   WPI; 1993-303469/38.
DR   N-PSDB; AAQ47881.
XX
PT   Gene coding for flavonoid-3',5'-hydroxylase of petunia petals - used to
PT   transform plants e.g. petunia, rose or tobacco to give bluer flower
PT   colour and altered pigment pattern.
XX
PS   Claim 11; Page 59; 82pp; Japanese.
XX
CC   Insertion of the sequences (AAQ47840-42) into plants such as rose,
CC   petunia, tobacco and carnation, using a suitable vector such as
CC   agrobacterium, give transformed plants which express the gene, resulting
CC   in petals with a bluer colour than normal, and/or pigmentation patterns
CC   which do not occur naturally. The sequences were amplified using primers
CC   (AAQ47843-70). Related single specific primers using a gene sequence
CC   coding for the haem-binding region of cytochrome P450 are shown in
CC   (AAQ47871-Q47903). (Updated on 25-MAR-2003 to correct PN field.) (Updated
CC   on 24-OCT-2003 to standardise OS field)
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 33;  DB 2;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGGGPRKCVG 11


RESULT 30
AAR40870
ID   AAR40870 standard; protein; 11 AA.
XX
AC   AAR40870;
XX
DT   24-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   28-MAR-1994  (first entry)
XX
DE   SSP for flavonoid-3',5'-hydroxylase gene product.
XX
KW   Flavonoid-3',5'-hydoxylase; transformation; plants; petunia; rose;
KW   tobacco; pigment alteration; blue; SSP; single specific primer; PCR;
KW   polymerase chain reaction; amplification; expression.
XX
OS   Petunia x hybrida.
XX
PN   WO9318155-A1.
XX
PD   16-SEP-1993.
XX
PF   20-NOV-1992;   92WO-JP001520.
XX
PR   02-MAR-1992;   92JP-00044963.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO CO LTD.
XX
PI   Kikuchi Y,  Kiyokawa S,  Shimada Y,  Ohbayashi M,  Shimada R;
PI   Okinaka Y;
XX
DR   WPI; 1993-303469/38.
DR   N-PSDB; AAQ47871.
XX
PT   Gene coding for flavonoid-3',5'-hydroxylase of petunia petals - used to
PT   transform plants e.g. petunia, rose or tobacco to give bluer flower
PT   colour and altered pigment pattern.
XX
PS   Claim 11; Page 55; 82pp; Japanese.
XX
CC   Insertion of the sequences (AAQ47840-42) into plants such as rose,
CC   petunia, tobacco and carnation, using a suitable vector such as
CC   agrobacterium, give transformed plants which express the gene, resulting
CC   in petals with a bluer colour than normal, and/or pigmentation patterns
CC   which do not occur naturally. The sequences were amplified using primers
CC   (AAQ47843-70). Related single specific primers using a gene sequence
CC   coding for the haem-binding region of cytochrome P450 are shown in
CC   (AAQ47871-Q47903). (Updated on 25-MAR-2003 to correct PN field.) (Updated
CC   on 24-OCT-2003 to standardise OS field)
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 33;  DB 2;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGSGRRICPG 11


RESULT 31
AAR40871
ID   AAR40871 standard; protein; 11 AA.
XX
AC   AAR40871;
XX
DT   24-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   28-MAR-1994  (first entry)
XX
DE   SSP for flavonoid-3',5'-hydroxylase gene product.
XX
KW   Flavonoid-3',5'-hydoxylase; transformation; plants; petunia; rose;
KW   tobacco; pigment alteration; blue; SSP; single specific primer; PCR;
KW   polymerase chain reaction; amplification; expression.
XX
OS   Petunia x hybrida.
XX
PN   WO9318155-A1.
XX
PD   16-SEP-1993.
XX
PF   20-NOV-1992;   92WO-JP001520.
XX
PR   02-MAR-1992;   92JP-00044963.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO CO LTD.
XX
PI   Kikuchi Y,  Kiyokawa S,  Shimada Y,  Ohbayashi M,  Shimada R;
PI   Okinaka Y;
XX
DR   WPI; 1993-303469/38.
DR   N-PSDB; AAQ47872.
XX
PT   Gene coding for flavonoid-3',5'-hydroxylase of petunia petals - used to
PT   transform plants e.g. petunia, rose or tobacco to give bluer flower
PT   colour and altered pigment pattern.
XX
PS   Claim 11; Page 55; 82pp; Japanese.
XX
CC   Insertion of the sequences (AAQ47840-42) into plants such as rose,
CC   petunia, tobacco and carnation, using a suitable vector such as
CC   agrobacterium, give transformed plants which express the gene, resulting
CC   in petals with a bluer colour than normal, and/or pigmentation patterns
CC   which do not occur naturally. The sequences were amplified using primers
CC   (AAQ47843-70). Related single specific primers using a gene sequence
CC   coding for the haem-binding region of cytochrome P450 are shown in
CC   (AAQ47871-Q47903). (Updated on 25-MAR-2003 to correct PN field.) (Updated
CC   on 24-OCT-2003 to standardise OS field)
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 33;  DB 2;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGAGRRICPG 11


RESULT 32
AAR41635
ID   AAR41635 standard; protein; 11 AA.
XX
AC   AAR41635;
XX
DT   24-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   28-MAR-1994  (first entry)
XX
DE   SSP for flavonoid-3',5'-hydroxylase gene product.
XX
KW   Flavonoid-3',5'-hydoxylase; transformation; plants; petunia; rose;
KW   tobacco; pigment alteration; blue; SSP; single specific primer; PCR;
KW   polymerase chain reaction; amplification; expression.
XX
OS   Petunia x hybrida.
XX
PN   WO9318155-A1.
XX
PD   16-SEP-1993.
XX
PF   20-NOV-1992;   92WO-JP001520.
XX
PR   02-MAR-1992;   92JP-00044963.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO CO LTD.
XX
PI   Kikuchi Y,  Kiyokawa S,  Shimada Y,  Ohbayashi M,  Shimada R;
PI   Okinaka Y;
XX
DR   WPI; 1993-303469/38.
DR   N-PSDB; AAQ47885.
XX
PT   Gene coding for flavonoid-3',5'-hydroxylase of petunia petals - used to
PT   transform plants e.g. petunia, rose or tobacco to give bluer flower
PT   colour and altered pigment pattern.
XX
PS   Claim 11; Page 61; 82pp; Japanese.
XX
CC   Insertion of the sequences (AAQ47840-42) into plants such as rose,
CC   petunia, tobacco and carnation, using a suitable vector such as
CC   agrobacterium, give transformed plants which express the gene, resulting
CC   in petals with a bluer colour than normal, and/or pigmentation patterns
CC   which do not occur naturally. The sequences were amplified using primers
CC   (AAQ47843-70). Related single specific primers using a gene sequence
CC   coding for the haem-binding region of cytochrome P450 are shown in
CC   (AAQ47871-Q47903). (Updated on 25-MAR-2003 to correct PN field.) (Updated
CC   on 24-OCT-2003 to standardise OS field)
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 33;  DB 2;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGGGRRICPG 11


RESULT 33
AAR40875
ID   AAR40875 standard; protein; 11 AA.
XX
AC   AAR40875;
XX
DT   24-OCT-2003  (revised)
DT   25-MAR-2003  (revised)
DT   28-MAR-1994  (first entry)
XX
DE   SSP for flavonoid-3',5'-hydroxylase gene.
XX
KW   Flavonoid-3',5'-hydoxylase; transformation; plants; petunia; rose;
KW   tobacco; pigment alteration; blue; SSP; single specific primer; PCR;
KW   polymerase chain reaction; amplification; expression; ss.
XX
OS   Petunia x hybrida.
XX
PN   WO9318155-A1.
XX
PD   16-SEP-1993.
XX
PF   20-NOV-1992;   92WO-JP001520.
XX
PR   02-MAR-1992;   92JP-00044963.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO CO LTD.
XX
PI   Kikuchi Y,  Kiyokawa S,  Shimada Y,  Ohbayashi M,  Shimada R;
PI   Okinaka Y;
XX
DR   WPI; 1993-303469/38.
DR   N-PSDB; AAQ47876.
XX
PT   Gene coding for flavonoid-3',5'-hydroxylase of petunia petals - used to
PT   transform plants e.g. petunia, rose or tobacco to give bluer flower
PT   colour and altered pigment pattern.
XX
PS   Claim 11; Page 57; 82pp; Japanese.
XX
CC   Insertion of the sequences (AAQ47840-42) into plants such as rose,
CC   petunia, tobacco and carnation, using a suitable vector such as
CC   agrobacterium, give transformed plants which express the gene, resulting
CC   in petals with a bluer colour than normal, and/or pigmentation patterns
CC   which do not occur naturally. The sequences were amplified using primers
CC   (AAQ47843-70). Related single specific primers using a gene sequence
CC   coding for the haem-binding region of cytochrome P450 are shown in
CC   (AAQ47871-Q47903). (Updated on 25-MAR-2003 to correct PN field.) (Updated
CC   on 24-OCT-2003 to standardise OS field)
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 33;  DB 2;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGAGRRVCPG 11


RESULT 34
AAU80018
ID   AAU80018 standard; peptide; 11 AA.
XX
AC   AAU80018;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Conserved haem-binding domain of P450 proteins.
XX
KW   Plant; aspen; phenylpropanoid pathway; agronomic; lignin; paper;
KW   4-coumarate-CoA ligase; 4CL; coniferyl aldehyde 5-hydroxylase; CA1d5H;
KW   S-adenosyl-L-methionine-dependent; SAM; AldOMT: transgenic; grass;
KW   5-hydroxyconiferaldehyde O-methyltransferase; cellulose; pulp;
KW   coniferyl alcohol dehydrogenase; CAD; sinapyl alcohol dehydrogenase; SAD;
KW   syringyl; guaiacyl; agriculture; haem-binding domain; P450.
XX
OS   Synthetic.
```

```
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 4
FT                   /label= unknown
FT   Misc-difference 6
FT                   /label= unknown
FT   Misc-difference 7
FT                   /label= unknown
FT   Misc-difference 8
FT                   /label= unknown
FT   Misc-difference 10
FT                   /label= unknown
XX
PN   WO200220717-A2.
XX
PD   14-MAR-2002.
XX
PF   05-SEP-2001; 2001WO-US027445.
XX
PR   05-SEP-2000; 2000US-0230086P.
XX
PA   (UNMT ) UNIV MICHIGAN TECHNOLOGICAL.
XX
PI   Chiang VLC,  Li L;
XX
DR   WPI; 2002-351773/38.
XX
PT   Genetically transforming plant with multiple genes from phenylpropanoid
PT   pathways, comprises incorporating number of genes into the genome of the
PT   plant, to produce plants displaying altered agronomic traits.
XX
PS   Disclosure; Page 6; 95pp; English.
XX
CC   The invention relates to a method of genetically transforming a plant
CC   simultaneously with multiple genes from the phenylpropanoid pathways,
CC   comprising incorporating into the genome of the plant, a number of genes,
CC   their substantially similar fragments or their combinations, to produce
CC   plants displaying altered agronomic traits. The genes are selected from 4
CC   -coumarate-CoA ligase (4CL), coniferyl aldehyde 5-hydroxylase (CA1d5H), S
CC   -adenosyl-L-methionine (SAM)-dependent 5-hydroxyconiferaldehyde O-
CC   methyltransferase (AldOMT), coniferyl alcohol dehydrogenase (CAD) and
CC   sinapyl alcohol dehydrogenase (SAD). The method is useful for the
CC   transformation of plant tissue for the alteration of lignin monomer
CC   composition, increased syringyl/guaiacyl (S/G) lignin ratio and increased
CC   cellulose content and transgenic plants resulting from such
CC   transformations. This is an improved method to simultaneously control the
CC   lignin quantity, lignin compositions, and cellulose contents in plants,
CC   and is applicable to all plant species that are susceptible to the
CC   transfer of genetic information by Agrobacterium or other gene delivery
CC   system. The method is of particular value to paper and pulp industries
CC   because lignin containing higher syringyl monomer content is more
CC   susceptible to chemical delignification. Woody plants transformed with
CC   DNA constructs offer a significant advantage in the delignification
CC   process over conventional paper feedstocks. Similarly, modification of
CC   the lignin composition in grasses by insertion and expression of
CC   heterologous SAD gene offers a unique method for increasing the
CC   digestibility of grasses and is of significant potential economic benefit
CC   to the farm and agricultural industries. The present sequence represents
CC   the amino acid sequence of the conserved haem-binding domain of P450
CC   proteins, which has homology to CAld5H as described in the method of the
CC   invention
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 33;  DB 5;  Length 11;
  Best Local Similarity   90.0%;  Pred. No. 3.9e+02;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |:||||||||
Db          2 FGXGXXXCXG 11
```

```
RESULT 35
ADO58606
ID   ADO58606 standard; peptide; 11 AA.
XX
AC   ADO58606;
XX
DT   15-JUL-2004  (first entry)
XX
DE   Pepper cytochrome P450 (pepCYP) associated peptide.
XX
KW   Pepper; cytochrome P450; PepCYP; incompatible interaction.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1. .11
FT                   /note= "All Xaa residues are unknown"
XX
PN   KR2001081198-A.
XX
PD   29-AUG-2001.
XX
PF   11-FEB-2000; 2000KR-00006345.
XX
PR   11-FEB-2000; 2000KR-00006345.
XX
PA   (KOKU-) KOREA KUMHO PETROCHEMICAL CO LTD.
XX
PI   Kim YS,  Ko MG,  Oh BJ;
XX
DR   WPI; 2002-128887/17.
XX
PT   Pepper cytochrom p450 gene(pepcyp) highly expressed during incompatible
PT   interaction between colletotrichum gloeosporioides and peppers.
XX
PS   Disclosure; Page 2; 16pp; Korean.
XX
CC   The present invention relates to the isolation of a pepper (Capsicum
CC   annuum) cytochrome P450 gene (PepCYP), which is highly expressed by
CC   incompatible interaction between Colletotrichum gloeosporioides and
CC   peppers. The present peptide sequence of unknown function is given in the
CC   specification of the present invention.
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 33;  DB 5;  Length 11;
  Best Local Similarity   80.0%;  Pred. No. 3.9e+02;
  Matches    8;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |:|||:||||
Db          2 FGXGXRXCXG 11


RESULT 36
AAE30014
ID   AAE30014 standard; peptide; 11 AA.
XX
AC   AAE30014;
XX
DT   24-FEB-2003  (first entry)
XX
DE   Cytochrome P450 protein conserved domain #3.
XX
KW   Tobacco; plant; cytochrome P450; disease resistance; biosynthesis; lipid;
KW   hormone; insect attractant; secondary metabolite; herbicide; pesticide;
KW   pollutant; deterrent; insect predator; flavouring; antibacterial;
KW   fungicide; virucide; insecticide.
XX
OS   Nicotiana tabacum.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 4
```

```
FT                   /label= Unknown
FT                   /note= "Xaa can be any amino acid"
FT   Misc-difference 8
FT                   /label= Unknown
FT                   /note= "Xaa can be any amino acid"
XX
PN   WO200272758-A2.
XX
PD   19-SEP-2002.
XX
PF   08-MAR-2002; 2002WO-US006912.
XX
PR   09-MAR-2001; 2001US-0274421P.
PR   13-MAR-2001; 2001US-0275597P.
XX
PA   (KENT ) UNIV KENTUCKY RES FOUND.
XX
PI   Chappell J,  Ralston LF;
XX
DR   WPI; 2003-018735/01.
XX
PT   New cytochrome P450 polypeptide useful for the biosynthesis of hormones,
PT   lipids and secondary metabolites, and in increasing the resistance of
PT   plants against harmful exogenous chemicals, e.g. herbicides, pesticides
PT   or pollutants.
XX
PS   Disclosure; Fig 4A; 78pp; English.
XX
CC   The invention relates to a new isolated cytochrome P450 polypeptide. The
CC   invention is useful for conferring disease resistance on a plant or plant
CC   component conferring disease resistance on a plant or plant component.
CC   The cytochrome p450 polypeptide is useful in the biosynthesis of
CC   hormones, lipids, and secondary metabolites, and may also help plants
CC   tolerate potentially harmful exogenous chemicals such as herbicides,
CC   pesticides and pollutants. In addition, the polypeptide is useful in the
CC   chemical defense of the plants against insects, as well as against
CC   bacterial, viral, or fungal infection. The polypeptide may also be used
CC   in the production of insect attractants and deterrents, which may also
CC   deter insect pests or attract insect predators, and in generating
CC   flavourings. The present sequence is Nicotiana tabacum cytochrome P450
CC   protein conserved domain
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 33;  DB 6;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 3.9e+02;
  Matches    6;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |:||::||:|
Db          2 FGXGRRXCPG 11


RESULT 37
AAR30868
ID   AAR30868 standard; protein; 13 AA.
XX
AC   AAR30868;
XX
DT   25-MAR-2003  (revised)
DT   12-MAY-1993  (first entry)
XX
DE   Consensus haem-binding domain of microsomal cyt-P450.
XX
KW   Dihydrokaempferol; flavonoid; pigmentation; colour; amplification;
KW   cytochrome P450; advocado; petunia; ss.
XX
OS   Synthetic.
XX
PN   EP522880-A2.
XX
PD   13-JAN-1993.
XX
PF   10-JUL-1992;   92EP-00306379.
XX
PR   11-JUL-1991;   91AU-00007173.
PR   17-FEB-1992;   92AU-00000923.
XX
PA   (ITFL-) INT FLOWER DEV PTY LTD.
XX
PI   Holton TA,  Cornish EC,  Kovacic F,  Tanaka Y,  Lester DR;
XX
DR   WPI; 1993-010688/02.
XX
PT   Nucleic acid sequence encoding a di:hydro:kaempferol-hydroxylating enzyme
PT   - e.g. cytochrome P450 introduced into transgenic plants for controlling
PT   flavonoid pigmentation in plants and organisms.
XX
PS   Disclosure; Page 13; 66pp; English.
XX
CC   The consensus sequence of haem-binding domains for microsomal cytochrome
CC   P450 homologues was used to design synthetic PCR primers. The primers may
CC   be used in PCR for amplification of petal cytochrome P450 homologues. See
CC   also AAR30868-75. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 13 AA;

  Query Match             100.0%;  Score 33;  DB 2;  Length 13;
  Best Local Similarity   60.0%;  Pred. No. 4.5e+02;
  Matches    6;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::||:||:|
Db          4 FGAGXRXCLG 13


RESULT 38
AAY09192
ID   AAY09192 standard; peptide; 13 AA.
XX
AC   AAY09192;
XX
DT   20-JUL-1999  (first entry)
XX
DE   Soybean cytochrome P450 enzyme conserved peptide sequence.
XX
KW   Soybean; Solanaceae crop plant; cytochrome P450; transgenic plant;
KW   enzyme; phenylurea herbicide; herbicide resistance.
XX
OS   Glycine max.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1. .13
FT                   /note= "Xaa is any amino acid"
XX
PN   WO9919493-A2.
XX
PD   22-APR-1999.
XX
PF   05-OCT-1998;   98WO-US020807.
XX
PR   10-OCT-1997;   97US-00948564.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Siminszky B,  Dewey RE,  Corbin FT;
XX
DR   WPI; 1999-302532/25.
XX
PT   DNA encoding soybean cytochrome P450 enzymes.
XX
PS   Example 2; Page 26; 93pp; English.
XX
CC   The invention provides new DNA molecules (AAX60773-X60781) encoding
CC   soybean cytochrome P450 enzymes (AAY09183-Y09191) respectively. The DNA
CC   encoding the cytochrome P450 enzymes is useful for transformation of
```

```
CC   Solanaceae crop plants. Transgenic plants comprising DNA constructs
CC   having the P450 encoding nucleic acid sequences are resistant to
CC   phenylurea herbicides. The transgenic plants have increased resistance to
CC   phenylurea herbicides compared to wild-type plants of the same species.
CC   The plant crops, e.g. turfgrass, tobacco, potato, tomato, corn, rice,
CC   cotton, soybean, rape, wheat, oats, barley or rice are particularly
CC   resistant to fluometuron, linuron, chlortoluron or diuron
XX
SQ   Sequence 13 AA;

  Query Match           100.0%;  Score 33;  DB 2;  Length 13;
  Best Local Similarity  80.0%;  Pred. No. 4.5e+02;
  Matches    8;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |:|||:||||
Db          4 FGXGXRXCXG 13


RESULT 39
AAB11394
ID   AAB11394 standard; protein; 13 AA.
XX
AC   AAB11394;
XX
DT   12-SEP-2003  (revised)
DT   22-FEB-2001  (first entry)
XX
DE   Gerbera flavone synthase FNSII haem-binding peptide fragment.
XX
KW   Gerbera; transgenic plant; flavone synthase II; FNSII; anticancer;
KW   immunomodulator; naringenin; apigenin; ornamental plant; flower colour;
KW   pharmaceutical; cancer; treatment.
XX
OS   Gerbera hybrid cultivar.
XX
PN   DE19918365-A1.
XX
PD   26-OCT-2000.
XX
PF   22-APR-1999;   99DE-01018365.
XX
PR   22-APR-1999;   99DE-01018365.
XX
PA   (MART/) MARTENS S.
PA   (FORK/) FORKMANN G.
XX
PI   Martens S,  Forkmann G;
XX
DR   WPI; 2000-648348/63.
XX
PT   New nucleic acid encoding flavone synthase II, useful e.g. for producing
PT   transgenic plants with altered flower color or flavone content.
XX
PS   Disclosure; Fig 4A; 40pp; German.
XX
CC   This invention describes a novel nucleic acid (I) that encodes flavone
CC   synthase II (FNSII) which has anticancer and immunomodulatory activity.
CC   FNSII catalyses conversion of naringenin to apigenin. (I) is used to
CC   produce transgenic ornamental plants that have targeted alterations in
CC   flower color, also altered content/distribution of flavones in leaves,
CC   flowers and other tissues, e.g. increased resistance properties or
CC   symbiotic capacity. FNSII expressed by (I) is used in synthesis of
CC   flavones that are useful as pharmaceuticals, e.g. in cancer treatment, or
CC   as biologically active substances, e.g. to improve the immune defence
CC   system. Oligonucleotide fragments of (I) are used as probes and primers,
CC   or as antisense or ribozyme agents for regulating expression of (II).
CC   (Updated on 12-SEP-2003 to standardise OS field)
XX
SQ   Sequence 13 AA;

  Query Match           100.0%;  Score 33;  DB 3;  Length 13;
  Best Local Similarity  70.0%;  Pred. No. 4.5e+02;
  Matches    7;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |:||::||||
Db          3 FGXGRRXCXG 12


RESULT 40
AAG98132
ID   AAG98132 standard; peptide; 14 AA.
XX
AC   AAG98132;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human SNP associated peptide SEQ ID NO. 774.
XX
KW   Human; single nucleotide polymorphism; SNP; angiopoietin;
KW   4-hydroxybutyrate; dehydrogenase; protein therapy;
KW   adenosine triphosphate-dependent RNA helicase;
KW   major histocompatibility complex Class I histocompatibility antigen; MHC;
KW   phosphoglycerate kinase; immunosuppressive; immunostimulatory;
KW   antirheumatic; antisclerotic; antidiabetic; antiinflammatory; cytostatic;
KW   antileukemic; neuroprotective; antimicrobial; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200148245-A2.
XX
PD   05-JUL-2001.
XX
PF   27-DEC-2000; 2000WO-US035346.
XX
PR   27-DEC-1999;   99US-00472688.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-418297/44.
XX
PT   Polymorphic nucleic acids encoding e.g. angiopoietin, dehydrogenase,
PT   adenosine triphosphate-dependent RNA helicase and/or phosphoglycerate
PT   kinase, useful for diagnosing and treating, e.g. cancer, autoimmune
PT   diseases and infections.
XX
PS   Disclosure; Page 461; 484pp; English.
XX
CC   The invention relates to nucleic acids (AAH79386-AAH80036) encoding
CC   polymorphic variants of proteins (AAG98010-AAG98238) related to
CC   angiopoietin, 4-hydroxybutyrate, dehydrogenase, adenosine triphosphate
CC   (ATP)-dependent RNA helicase, major histocompatibility complex (MHC)
CC   Class I histocompatibility antigen and/or phosphoglycerate kinase. These
CC   nucleic acid single nucleotide polymorphisms (SNPs) and the encoded
CC   proteins have potential immunosuppressive, immunostimulatory,
CC   antirheumatic, antisclerotic, antidiabetic, antiinflammatory, cytostatic,
CC   antileukemic, neuroprotective and antimicrobial activity and may be
CC   useful in gene/protein therapy, vaccines, modulation of the expression
CC   and activity of proteins related to angiopoietin, 4-hydroxybutyrate,
CC   dehydrogenase, adenosine triphosphate (ATP)-dependent RNA helicase, major
CC   histocompatibility complex (MHC) Class I histocompatibility antigen
CC   and/or phosphoglycerate kinase. Disorders that may be prevented,
CC   diagnosed and/or treated by the above methods include multifactorial
CC   diseases with a genetic component, such as autoimmune diseases (e.g.
CC   rheumatoid arthritis, multiple sclerosis, diabetes, systemic lupus
CC   erythematosus and Grave's disease), inflammation, cancer (e.g. cancers of
CC   the bladder, brain, breast, colon and kidney, leukemia), diseases of the
CC   nervous system, an infection of pathogenic organisms. They may also be
CC   used to alter phenotypic traits such as longevity, appearance, strength,
CC   speed and endurance
XX
SQ   Sequence 14 AA;
```

```
  Query Match             100.0%;  Score 33;  DB 4;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 4.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FHYGVLACEG 11


RESULT 41
AAP20076
ID   AAP20076 standard; protein; 15 AA.
XX
AC   AAP20076;
XX
DT   25-MAR-2003  (revised)
DT   01-DEC-1992  (first entry)
XX
DE   Synthetic peptide specific antigenic determinant region i.
XX
KW   Antigen; vaccine; diagnostic; therapeutic antibody.
XX
OS   Synthetic.
XX
PN   EP44710-A.
XX
PD   27-JAN-1982.
XX
PF   17-JUL-1980;   80US-00169758.
XX
PR   17-JUL-1980;   80US-00169758.
PR   30-OCT-1980;   80US-00202431.
PR   27-MAR-1981;   81US-00248059.
XX
PA   (SCRI ) SCRIPPS CLINIC & RES FOUND.
PA   (SCHR-) SCHRIPPS CLINIC RES.
XX
PI   Lerner RA,  Green N,  Sutcliffe JG,  Shinnick TM;
XX
DR   WPI; 1982-08369E/05.
XX
PT   Synthetic specific antigenic determinants - comprising peptides with
PT   amino acid sequence determined from gene DNA sequence.
XX
PS   Claim 14; Page 75; 93pp; English.
XX
CC   The peptide is a synthetic peptide specific antigenic determinant region,
CC   it is synthesised based on the sequence of a specific antigenic
CC   determinant of a desired natural genome. It can be used in the prodn. of
CC   antigens which can be used to produce vaccines, diagnostic or therapeutic
CC   antibodies etc. The antigens produced are highly specific and free of
CC   undesirable impurities. See also AAP20068-P20094. (Updated on 25-MAR-2003
CC   to correct PA field.)
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 33;  DB 1;  Length 15;
  Best Local Similarity   40.0%;  Pred. No. 5.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          4 FLGGWWVCLG 13


RESULT 42
AAP50146
ID   AAP50146 standard; peptide; 15 AA.
XX
AC   AAP50146;
XX
DT   25-MAR-2003  (revised)
DT   16-AUG-2002  (revised)
DT   07-OCT-1991  (first entry)
XX
DE   Sequence of synthetic peptide used in vaccine against hepatitis B virus
DE   (HBV).
XX
KW   Vaccine; immunogen; antigen; epitope; diagnosis.
XX
OS   Hepatitis B virus.
XX
PN   EP155146-A.
XX
PD   18-SEP-1985.
XX
PF   07-MAR-1985;   85EP-00301588.
XX
PR   09-MAR-1984;   84US-00587983.
PR   09-MAR-1984;   84US-00588122.
XX
PA   (SCRI ) SCRIPPS CLINIC & RES FOUND.
XX
PI   Milich DR,  Chisari FV;
XX
DR   WPI; 1985-237984/39.
XX
PT   New synthetic polypeptide(s) - useful in vaccines against hepatitis B
PT   virus and for diagnostic tests.
XX
PS   Claim 1a; Page 61; 76pp; English.
XX
CC   The inventors claim a vaccine against HBV which comprises one of AAP50146
CC   -P50151 together with one of AAP50152-P50156 and a diluent. They also
CC   claim a vaccine which comprises a synthetic polypeptide having an AA
CC   residue sequence immunologically corresp. to portions of a sequence of
CC   the virus surface antigen (a) from posns. 38-52, and (b) from posns. 110-
CC   137, from the amino terminus, together with a diluent. Other portions of
CC   sequences may be similarly used, (see AAP50158, AAP50160, AAP50162).
CC   (Updated on 16-AUG-2002 to add missing OS field.) (Updated on 25-MAR-2003
CC   to correct PA field.)
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 33;  DB 1;  Length 15;
  Best Local Similarity   40.0%;  Pred. No. 5.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          4 FLGGTTVCLG 13


RESULT 43
AAP60758
ID   AAP60758 standard; protein; 15 AA.
XX
AC   AAP60758;
XX
DT   25-MAR-2003  (revised)
DT   01-JAN-1980  (first entry)
XX
DE   Labeled synthetic peptide used to recognise hepatitis B virus surface
DE   antigen or antibody.
XX
KW   Hepatitis B virus; surface antigen; antibody; radiolabel; enzyme; label;
KW   disease diagnosis; ss.
XX
OS   Synthetic.
XX
PN   US4591552-A.
XX
PD   27-MAY-1986.
XX
PF   29-SEP-1982;   82US-00426309.
XX
```

PR 29-SEP-1982; 82US-00426309.
XX
PA (NYBL-) NEW YORK BLOOD CENTER INC.
XX
PI Neurath AR;
XX
DR WPI; 1986-155507/24.
XX
PT Detection of Hepatitis B surface antigen or antibody - using labelled
PT synthetic peptide.
XX
PS Claim 7; Page 23; 15pp; English.
XX
CC This radio-labeled or enzyme-labeled peptide is used in the detection of
CC hepatitis B virus surface antigen or antibody. The method involves
CC contacting the sample with the labeled peptide which is normally
CC recognized and bound by the antibody suspected of being present in the
CC sample. The contacted sample is incubated and washed. Staphylococci
CC Protein-A are then added to the contacted sample with the resultant mass
CC being incubated and centrifuged. The radioactivity or enzymatic activity
CC of the resultant pelleted bacteria is then determined. See also AAP60750-
CC 7, AAP60759-61. (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 33; DB 1; Length 15;
Best Local Similarity 40.0%; Pred. No. 5.1e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|::|:::|:|
Db 4 FLGGTTVCLG 13

RESULT 44
AAR24422
ID AAR24422 standard; peptide; 15 AA.
XX
AC AAR24422;
XX
DT 25-MAR-2003 (revised)
DT 21-NOV-1992 (first entry)
XX
DE Sequence of T helper determinant in the S region of HBsAg from amino acid
DE residue numbers 38-52.
XX
KW Vaccine; AIDS; HIV-1; carrier peptide.
XX
OS Homo sapiens.
XX
PN WO9208491-A1.
XX
PD 29-MAY-1992.
XX
PF 19-NOV-1991; 91WO-US008653.
XX
PR 20-NOV-1990; 90US-00616247.
XX
PA (TANO-) TANOX BIOSYSTEMS INC.
XX
PI Chang TW, Fung MSC;
XX
DR WPI; 1992-199955/24.
XX
PT Vaccines comprising anti-idiotype antibody conjugates - induce prodn. of
PT neutralising antibodies against HIV-1 for immunisation against HIV
PT infection and AIDS.
XX
PS Claim 9; Page 25 and page 16-17; 29pp; English.
XX
CC The invention includes epitope-directed immunization with a vaccine in
CC which an anti-idiotype antibody is conjugated to a carrier, which can be
CC either a protein or its derived T helper peptide. The carrier is one
CC against which the vaccine recipient has previously immunized or otherwise
CC previously exposed, or which enhances the immune response against the
CC anti-idiotype antibody. One exemplary anti-idiotype antibody which
CC induces antibodies against the PND is AB19-4. Where the anti-idiotype
CC induces Ab3 against HIV-1, the carrier preferably is HBsAg or HIV-1 p24,
CC or a peptide of either HBsAg or HIV-1 p24 including a T helper
CC determinant. "PND" = the principal neutralizing determinant ("PND") of
CC gp120. It is known that many individuals in the high risk groups for HIV-
CC 1 infection are also at high risk for hepatitis B virus infection. Many
CC people in these groups have already been vaccinated with the hepatitis B
CC virus vaccine using hepatitis B surface antigen (HBsAg). Thus, HBsAg (and
CC related T helper peptides) is a preferred carrier for conjugation with an
CC anti-idiotype antibody which induces production of HIV-1 neutralizing
CC antibodies. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 33; DB 2; Length 15;
Best Local Similarity 40.0%; Pred. No. 5.1e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|::|:::|:|
Db 4 FLGGTTVCLG 13

RESULT 45
AAR82090
ID AAR82090 standard; peptide; 15 AA.
XX
AC AAR82090;
XX
DT 25-MAR-2003 (revised)
DT 03-APR-1996 (first entry)
XX
DE Hepatitis B virus major surface antigen 38-52 T cell epitope.
XX
KW Retro-inverso modified; T cell epitope; analogue; vaccine; B cell;
KW immunisation; HBV; major surface antigen; residues 38-52.
XX
OS Hepatitis B virus.
XX
PN WO9523166-A1.
XX
PD 31-AUG-1995.
XX
PF 24-FEB-1995; 95WO-AU000090.
XX
PR 25-FEB-1994; 94AU-00004119.
XX
PA (DEAK-) DEAKIN RES LTD.
XX
PI Comis A, Tyler MI, Fischer P;
XX
DR WPI; 1995-311503/40.
XX
PT Synthetic peptide T cell epitope analogue (retro-)inverso modified - used
PT conjugated with a B cell epitope in a vaccine tailored to a specific
PT condition, e.g. polio, hepatitis B, etc.
XX
PS Claim 3; Page 34; 51pp; English.
XX
CC A vaccine comprising a retro-inverso modified, native T cell epitope i.e.
CC AAR82073-R82094, conjugated to a B-cell epitope (i.e. the malaria
CC immunodominant B-cell epitope (AAR82095), and epitopes based on
CC poliomyelitis, tetanus or hepatitis B proteins etc.) is tailored to a
CC condition of interest (depending on the epitopes used), and used to
CC immunise a host against the specific condition. (Updated on 25-MAR-2003
CC to correct PI field.)
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 33; DB 2; Length 15;

```
  Best Local Similarity   40.0%;  Pred. No. 5.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          4 FLGGTTVCLG 13


RESULT 46
AAB19305
ID   AAB19305 standard; peptide; 15 AA.
XX
AC   AAB19305;
XX
DT   19-FEB-2001  (first entry)
XX
DE   Immunogenic peptide useful for generating antigen specific T cells.
XX
KW   Immunogenic peptide; antigen specific T cell; cell proliferation; cancer;
KW   inflammation; autoimmune disease; dermatological disorder;
KW   neurodegenerative disorder; atherosclerosis; rheumatoid arthritis;
KW   osteoporosis; chronic ulcer; psoriasis; cardiovascular disease;
KW   infectious disease.
XX
OS   Hepatitis A virus.
XX
PN   WO200057920-A2.
XX
PD   05-OCT-2000.
XX
PF   30-MAR-2000; 2000WO-GB001225.
XX
PR   30-MAR-1999;   99GB-00007366.
XX
PA   (MEDI-) MEDICAL RES COUNCIL.
XX
PI   Fisher A;
XX
DR   WPI; 2000-647208/62.
XX
PT   Regulating expression of transgenes encoding enzymes or hormones in
PT   mammals, by transforming expandable population of cells with desired
PT   transgene and regulating cell proliferation by administration of an
PT   agent.
XX
PS   Disclosure; Page 22; 37pp; English.
XX
CC   The present sequence represents an immunogenic peptide, derived from the
CC   S antigen, which is useful for generating antigen specific T cells. It
CC   may be used in the course of the invention. The specification describes
CC   the use of an agent which regulates cell proliferation for modulating the
CC   levels of production of a gene product of interest in a host organism.
CC   The method comprises transforming an expandable population of cells with
CC   a transgene encoding the gene product, expressing the gene in the host
CC   and regulating proliferation of the population of cells by administration
CC   of the agent. The method is useful for regulating cell proliferation. The
CC   method is also useful for delivering one or more transgenes useful in the
CC   treatment of cancer, inflammation, autoimmune disease, dermatological
CC   disorder, neurodegenerative disorder, atherosclerosis, rheumatoid
CC   arthritis, osteoporosis, chronic ulcers, psoriasis, cardiovascular
CC   diseases and infectious diseases
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 33;  DB 3;  Length 15;
  Best Local Similarity   40.0%;  Pred. No. 5.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          4 FLGGTTVCLG 13


RESULT 47
AAU91319
ID   AAU91319 standard; peptide; 15 AA.
XX
AC   AAU91319;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Cytochrome P450 haeme binding motif.
XX
KW   Thymus expressed cytochrome; arachidonic acid metabolism;
KW   autoimmune disorder; Addison's disease; haemolytic anaemia;
KW   rheumatoid arthritis; dermatitis; allergic encephalomyelitis;
KW   glomerulonephritis; Goodpasture's Syndrome; Grave's disease;
KW   multiple sclerosis; myasthenia gravis; neuritis; organ rejection;
KW   systemic lupus erythematosus; insulin dependent diabetes mellitus;
KW   autoimmune inflammatory eye disease; graft versus-host disease;
KW   inflammatory disease; infection; septic shock; sepsis; SIRS;
KW   systemic inflammatory response syndrome; ischaemia reperfusion injury;
KW   arthritis; inflammatory bowel disease; Crohn's disease; thymoma; P450TEC;
KW   immune system disease; haeme binding motif.
XX
OS   Unidentified.
XX
PN   WO200181585-A2.
XX
PD   01-NOV-2001.
XX
PF   20-APR-2001; 2001WO-CA000547.
XX
PR   20-APR-2000; 2000US-0198617P.
PR   01-JUN-2000; 2000US-0208785P.
XX
PA   (CYTO-) CYTOCHROMA INC.
XX
PI   Jones G,  Petkovich PM,  White JA,  Ramshaw HA,  Stangle WA;
XX
DR   WPI; 2002-291654/33.
XX
PT   New human cytochrome P450TEC, thymus expressed cytochrome polypeptides
PT   for diagnosing, treating autoimmune disorders e.g. Addison's disease,
PT   dermatitis, rheumatoid arthritis and identifying modulators.
XX
PS   Example 1; Page 65; 154pp; English.
XX
CC   The invention relates to an isolated human cytochrome P450TEC (thymus
CC   expressed cytochrome) polypeptide or the encoded sequence included in
CC   ATCC Deposit No.PTA-1785. Also included are the polynucleotide encoding
CC   P450TEC, its allelic variants, complement, or species homologue, where
CC   the nucleic acid can modulate, or encodes a peptide that modulates
CC   arachidonic acid metabolism and hydroxylation of arachidonic acid.
CC   P450TEC is useful for treating disease or condition related to P450TEC
CC   and for identifying an activity in a biological assay, by expressing
CC   P450TEC in a cell, isolating the biological fraction, detecting an
CC   activity in a biological assay and identifying the protein in the
CC   supernatant having the activity. The nucleic acid is useful for diagnosis
CC   of P450TEC-related disease or condition, in particular autoimmune
CC   disease, a disease related to hydroxylation pathway of arachidonic acid
CC   or inflammatory response of a patient or a predisposition to the disease
CC   by detecting a polymorphism in P450TEC gene. P450TEC is useful for
CC   identifying modulators of P450TEC activity. An anti-P450TEC antibody is
CC   useful for quantifying the level of P450TEC in a sample and for treating
CC   disease or condition related to P450TEC. P450TEC polynucleotides or
CC   polypeptides are useful as marker or detector of a particular immune
CC   system disease or disorder and in treating or detecting autoimmune
CC   disorders including Addison's disease, haemolytic anaemia, rheumatoid
CC   arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis,
CC   Goodpasture's Syndrome, Grave's disease, multiple sclerosis, myasthenia
CC   gravis, neuritis, systemic lupus erythematosus, insulin dependent
CC   diabetes mellitus, autoimmune inflammatory eye disease and organ
CC   rejection or graft versus-host disease. P450TEC is useful for diagnosis
CC   or analysis of disease states involving P450TEC or to monitor progress of
CC   therapy. P450TEC antagonists are useful to treat disorders such as acute
```

CC and chronic inflammatory diseases, including inflammation associated with
CC infection (e.g. septic shock, sepsis, systemic inflammatory response
CC syndrome (SIRS)), ischaemia reperfusion injury, arthritis, inflammatory
CC bowel disease, Crohn's disease and conditions of the thymus, such as
CC thymomas. The present sequence is a cytochrome haeme binding motif used
CC in segeunce similarity searches to identify nucleic acid sequences
CC encoding P450TEC
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 33; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
||||||||||
Db 2 FXXGXXXCXG 11

RESULT 48
ABG75617
ID ABG75617 standard; peptide; 15 AA.
XX
AC ABG75617;
XX
DT 16-APR-2003 (first entry)
XX
DE Arabidopsis thaliana cytochrome P450 protein peptide fragment #5.
XX
KW Arabidopsis; cytochrome P450; cP450; plant; parthenocarpic; fruit;
KW enlarged fruit; transgenic; vegetable size; grain size; soybean;
KW leaf size; flower size; agricultural yield; male-sterile plant;
KW moth orchid; maize.
XX
OS Arabidopsis thaliana.
OS Glycine max.
OS Pinus radiata.
OS Zea mays.
OS Phalaenopsis sp.
XX
FH Key Location/Qualifiers
FT Misc-difference 7
FT /label= Ala, Ser
FT Misc-difference 9
FT /label= Arg, Lys
FT Misc-difference 11
FT /label= Val, Ile, Ala
XX
PN US2002152495-A1.
XX
PD 17-OCT-2002.
XX
PF 09-JUL-1999; 99US-00349385.
XX
PR 15-JAN-1999; 99US-0115967P.
XX
PA (ITOT/) ITO T.
PA (FROM/) FROMM M.
PA (MEYE/) MEYEROWITZ E.
XX
PI Ito T, Fromm M, Meyerowitz E;
XX
DR WPI; 2003-198390/19.
XX
PT New polynucleotides encoding a cytochrome P450 polypeptide, useful for
PT producing plants with enlarged or parthenocarpic fruits, in increasing
PT plant tissue size to increase agricultural yields, or producing male-
PT sterile plants.
XX
PS Disclosure; Page 3; 31pp; English.
XX
CC This invention relates to an isolated polynucleotide encoding a
CC cytochrome P450 polypeptide which when expressed in a plant produces at
CC least one phenotype selected from parthenocarpic fruit and enlarged
CC fruit. The invention also discloses a recombinant construct comprising
CC the novel polynucleotide which when expressed in a plant produces a plant
CC with a parthenocarpic fruit or enlarged fruit. Also disclosed is a
CC transgenic plant which expresses a cytochrome P450 or comprising an
CC isolated polynucleotide encoding a cytochrome P450 polypeptide which when
CC expressed in a plant produces a plant with a parthenocarpic fruit or
CC enlarged fruit compared with a plant lacking the isolated polynucleotide.
CC The cytochrome P450 polypeptide and polynucleotide of the invention is
CC useful in the production of plants with seedless, enlarged, or
CC parthenocarpic fruits, including vegetable or grain size, leaf size or
CC flower size. The sequences may also be used in increasing plant tissue
CC size to increase agricultural yields of plants, to produce male-sterile
CC plants and to screen for compounds that control parthenocarpy or fruit
CC size in plants. The present sequence represents a cytochrome P450 peptide
CC of the invention, this sequence is highly conserved in cytochrome P450
CC proteins
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 33; DB 6; Length 15;
Best Local Similarity 70.0%; Pred. No. 5.1e+02;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|:|||:|||:|
Db 5 FGXGXRXCPG 14

RESULT 49
ADV23624
ID ADV23624 standard; peptide; 15 AA.
XX
AC ADV23624;
XX
DT 10-MAR-2005 (first entry)
XX
DE HBV immunogenic peptide #425.
XX
KW Vaccine; virucide; antigen; autoimmune disease; infection;
KW immune modulation; cancer; neoplasm; cytostatic; melanoma; lung tumor;
KW breast tumor; uterine cervix tumor; prostatic cancer; colon tumor;
KW pancreas tumor; stomach tumor; bladder tumor; kidney tumor;
KW hodgkin's lymphoma.
XX
OS Hepatitis B virus.
XX
PN WO2004108753-A1.
XX
PD 16-DEC-2004.
XX
PF 10-JUN-2004; 2004WO-AU000775.
XX
PR 10-JUN-2003; 2003AU-00902875.
PR 25-MAR-2004; 2004AU-00901589.
XX
PA (UYME) UNIV MELBOURNE.
XX
PI Kent SJ;
XX
DR WPI; 2005-031657/03.
XX
PT Use of at least one set of peptides in the preparation of a medicament
PT for modulating an immune response, and for treating cancer or yeast,
PT viral, bacterial, protozoal and mycoplasma infections.
XX
PS Disclosure; SEQ ID NO 2044; 645pp; English.
XX
CC The invention relates to the use of at least one set of peptides in the
CC preparation of a medicament for modulating an immune response, where
CC individual peptides of a respective set comprise different portions of an
CC amino acid sequence corresponding to a single polypeptide of interest and
CC display partial sequence identity or similarity to at least one other

CC peptide of the same set of peptides (i.e. they are overlapping). Also
CC included are an antigen-presenting cell which has been contacted with the
CC peptides above and thus presents the peptides, a population of such
CC antigen-presenting cells, a process for producing antigen-presenting
CC cells for modulating an immune response to a polypeptide of interest, a
CC method for producing antigen-specific lymphocytes, a composition
CC comprising at least one set of the peptides (and a carrier and/or
CC diluent), a method for modulating an immune response to a polypeptide of
CC interest comprising administering to a patient in need at least one set
CC of the peptides, a method for treatment and/or prophylaxis of a disease
CC or condition associated with the presence of a polypeptide of interest
CC and a composition of matter for modulating an immune response in a
CC subject to a target antigen. The polypeptide of interest is also a
CC disease- or condition-associated polypeptide that is a polypeptide
CC produced by a pathogenic organism or a cancer, and produced by a
CC pathogenic organism selected from yeast, viruses, bacteria, helminths,
CC protozoans and mycoplasmas. The disease- or condition-associated
CC polypeptide is produced by a cancer selected from melanoma, lung cancer,
CC breast cancer, cervical cancer, prostate cancer, colon cancer, pancreatic
CC cancer, stomach cancer, bladder cancer, kidney cancer, post transplant
CC lymphoproliferative disease (PTLD) or Hodgkin's Lymphoma. The uncultured
CC antigen-presenting cells or their precursors are useful in the
CC preparation of a medicament for the treatment of a disease or condition
CC in a subject, which disease or condition is associated with the presence
CC or aberrant expression of a target antigen, where the antigen-presenting
CC cells or their precursors have not been subjected to activating
CC conditions but have been contacted with an antigen that corresponds to
CC the target antigen to express a processed or modified form of the antigen
CC for presentation to the subject's immune system. The present sequence is
CC one of a set of overlapping immunogenic peptides derived from a Hepatitis
CC B virus protein.
XX
SQ Sequence 15 AA;

```
  Query Match             100.0%;  Score 33;  DB 9;  Length 15;
  Best Local Similarity   40.0%;  Pred. No. 5.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FLGGTTVCLG 11
```

RESULT 50
ADV23625
ID ADV23625 standard; peptide; 15 AA.
XX
AC ADV23625;
XX
DT 10-MAR-2005 (first entry)
XX
DE HBV immunogenic peptide #426.
XX
KW Vaccine; virucide; antigen; autoimmune disease; infection;
KW immune modulation; cancer; neoplasm; cytostatic; melanoma; lung tumor;
KW breast tumor; uterine cervix tumor; prostatic cancer; colon tumor;
KW pancreas tumor; stomach tumor; bladder tumor; kidney tumor;
KW hodgkin's lymphoma.
XX
OS Hepatitis B virus.
XX
PN WO2004108753-A1.
XX
PD 16-DEC-2004.
XX
PF 10-JUN-2004; 2004WO-AU000775.
XX
PR 10-JUN-2003; 2003AU-00902875.
PR 25-MAR-2004; 2004AU-00901589.
XX
PA (UYME) UNIV MELBOURNE.
XX
PI Kent SJ;

XX
DR WPI; 2005-031657/03.
XX
PT Use of at least one set of peptides in the preparation of a medicament
PT for modulating an immune response, and for treating cancer or yeast,
PT viral, bacterial, protozoal and mycoplasma infections.
XX
PS Disclosure; SEQ ID NO 2045; 645pp; English.
XX
CC The invention relates to the use of at least one set of peptides in the
CC preparation of a medicament for modulating an immune response, where
CC individual peptides of a respective set comprise different portions of an
CC amino acid sequence corresponding to a single polypeptide of interest and
CC display partial sequence identity or similarity to at least one other
CC peptide of the same set of peptides (i.e. they are overlapping). Also
CC included are an antigen-presenting cell which has been contacted with the
CC peptides above and thus presents the peptides, a population of such
CC antigen-presenting cells, a process for producing antigen-presenting
CC cells for modulating an immune response to a polypeptide of interest, a
CC method for producing antigen-specific lymphocytes, a composition
CC comprising at least one set of the peptides (and a carrier and/or
CC diluent), a method for modulating an immune response to a polypeptide of
CC interest comprising administering to a patient in need at least one set
CC of the peptides, a method for treatment and/or prophylaxis of a disease
CC or condition associated with the presence of a polypeptide of interest
CC and a composition of matter for modulating an immune response in a
CC subject to a target antigen. The polypeptide of interest is also a
CC disease- or condition-associated polypeptide that is a polypeptide
CC produced by a pathogenic organism or a cancer, and produced by a
CC pathogenic organism selected from yeast, viruses, bacteria, helminths,
CC protozoans and mycoplasmas. The disease- or condition-associated
CC polypeptide is produced by a cancer selected from melanoma, lung cancer,
CC breast cancer, cervical cancer, prostate cancer, colon cancer, pancreatic
CC cancer, stomach cancer, bladder cancer, kidney cancer, post transplant
CC lymphoproliferative disease (PTLD) or Hodgkin's Lymphoma. The uncultured
CC antigen-presenting cells or their precursors are useful in the
CC preparation of a medicament for the treatment of a disease or condition
CC in a subject, which disease or condition is associated with the presence
CC or aberrant expression of a target antigen, where the antigen-presenting
CC cells or their precursors have not been subjected to activating
CC conditions but have been contacted with an antigen that corresponds to
CC the target antigen to express a processed or modified form of the antigen
CC for presentation to the subject's immune system. The present sequence is
CC one of a set of overlapping immunogenic peptides derived from a Hepatitis
CC B virus protein.
XX
SQ Sequence 15 AA;

```
  Query Match             100.0%;  Score 33;  DB 9;  Length 15;
  Best Local Similarity   40.0%;  Pred. No. 5.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          3 FLGGAPTCPG 12
```

Search completed: March 8, 2006, 11:19:58
Job time : 194 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2006, 11:25:42 ; Search time 23 Seconds
(without alignments)
12.102 Million cell updates/sec

Title: US-10-751-235-14
Perfect score: 33
Sequence: 1 FXXGXXXCXG 10

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 161667 seqs, 27834885 residues

Total number of hits satisfying chosen parameters: 161667

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	100.0	10	7	US-11-166-609-24	Sequence 24, Appl
2	33	100.0	10	7	US-11-166-609-25	Sequence 25, Appl
3	33	100.0	52	7	US-11-205-109-43	Sequence 43, Appl
4	33	100.0	61	7	US-11-096-568A-25378	Sequence 25378, A
5	33	100.0	71	7	US-11-087-099-10778	Sequence 10778, A
6	33	100.0	73	7	US-11-116-881A-542	Sequence 542, App
7	33	100.0	73	7	US-11-116-881A-546	Sequence 546, App
8	33	100.0	75	6	US-10-467-657-8472	Sequence 8472, Ap
9	33	100.0	84	6	US-10-453-372-494	Sequence 494, App
10	33	100.0	86	7	US-11-116-881A-548	Sequence 548, App
11	33	100.0	99	7	US-11-096-568A-25377	Sequence 25377, A
12	33	100.0	101	7	US-11-116-881A-2299	Sequence 2299, Ap
13	33	100.0	103	7	US-11-072-512-3124	Sequence 3124, Ap
14	33	100.0	107	7	US-11-096-568A-26958	Sequence 26958, A
15	33	100.0	113	6	US-10-821-234-1478	Sequence 1478, Ap
16	33	100.0	117	7	US-11-087-099-6613	Sequence 6613, Ap
17	33	100.0	118	7	US-11-116-881A-464	Sequence 464, App
18	33	100.0	119	7	US-11-116-881A-460	Sequence 460, App
19	33	100.0	126	7	US-11-072-512-2610	Sequence 2610, Ap
20	33	100.0	128	7	US-11-166-609-14	Sequence 14, Appl
21	33	100.0	128	7	US-11-166-609-15	Sequence 15, Appl
22	33	100.0	129	7	US-11-116-881A-500	Sequence 500, App
23	33	100.0	129	7	US-11-116-881A-516	Sequence 516, App
24	33	100.0	129	7	US-11-116-881A-540	Sequence 540, App
25	33	100.0	134	7	US-11-143-980-68	Sequence 68, Appl
26	33	100.0	135	7	US-11-116-881A-456	Sequence 456, App
27	33	100.0	137	7	US-11-116-881A-520	Sequence 520, App
28	33	100.0	140	7	US-11-116-881A-462	Sequence 462, App
29	33	100.0	141	7	US-11-084-508-22	Sequence 22, Appl
30	33	100.0	151	7	US-11-116-881A-448	Sequence 448, App
31	33	100.0	151	7	US-11-116-881A-510	Sequence 510, App
32	33	100.0	155	7	US-11-096-568A-12468	Sequence 12468, A
33	33	100.0	156	7	US-11-096-568A-14864	Sequence 14864, A
34	33	100.0	158	7	US-11-096-568A-14863	Sequence 14863, A
35	33	100.0	162	7	US-11-116-881A-486	Sequence 486, App
36	33	100.0	164	7	US-11-116-881A-484	Sequence 484, App
37	33	100.0	175	7	US-11-116-881A-470	Sequence 470, App
38	33	100.0	176	7	US-11-120-308-106	Sequence 106, App
39	33	100.0	176	7	US-11-116-881A-522	Sequence 522, App
40	33	100.0	177	7	US-11-116-881A-502	Sequence 502, App
41	33	100.0	185	6	US-10-467-657-1228	Sequence 1228, Ap
42	33	100.0	187	7	US-11-116-881A-458	Sequence 458, App
43	33	100.0	194	7	US-11-096-568A-23881	Sequence 23881, A
44	33	100.0	197	7	US-11-096-568A-18243	Sequence 18243, A
45	33	100.0	199	5	US-09-995-493-12	Sequence 12, Appl
46	33	100.0	208	7	US-11-087-099-7405	Sequence 7405, Ap
47	33	100.0	208	7	US-11-087-099-11148	Sequence 11148, A
48	33	100.0	210	7	US-11-096-568A-23880	Sequence 23880, A
49	33	100.0	213	7	US-11-087-099-2635	Sequence 2635, Ap
50	33	100.0	213	7	US-11-087-099-11719	Sequence 11719, A
51	33	100.0	213	7	US-11-096-568A-20260	Sequence 20260, A
52	33	100.0	219	7	US-11-096-568A-24257	Sequence 24257, A
53	33	100.0	221	7	US-11-096-568A-27263	Sequence 27263, A
54	33	100.0	222	7	US-11-019-711-81	Sequence 81, Appl
55	33	100.0	222	7	US-11-087-099-6134	Sequence 6134, Ap
56	33	100.0	222	7	US-11-096-568A-26174	Sequence 26174, A
57	33	100.0	224	7	US-11-087-099-8664	Sequence 8664, Ap
58	33	100.0	225	7	US-11-096-568A-32197	Sequence 32197, A
59	33	100.0	226	7	US-11-148-012-2	Sequence 2, Appli
60	33	100.0	226	7	US-11-148-012-3	Sequence 3, Appli
61	33	100.0	229	7	US-11-096-568A-32196	Sequence 32196, A
62	33	100.0	233	7	US-11-096-568A-24646	Sequence 24646, A
63	33	100.0	234	7	US-11-096-568A-15121	Sequence 15121, A
64	33	100.0	237	7	US-11-096-568A-24645	Sequence 24645, A
65	33	100.0	237	7	US-11-096-568A-31128	Sequence 31128, A
66	33	100.0	240	7	US-11-019-711-76	Sequence 76, Appl
67	33	100.0	240	7	US-11-087-099-1831	Sequence 1831, Ap
68	33	100.0	241	7	US-11-019-711-22	Sequence 22, Appl
69	33	100.0	241	7	US-11-019-711-77	Sequence 77, Appl
70	33	100.0	241	7	US-11-019-711-78	Sequence 78, Appl
71	33	100.0	241	7	US-11-234-786-114	Sequence 114, App
72	33	100.0	242	7	US-11-096-568A-31127	Sequence 31127, A
73	33	100.0	247	7	US-11-019-711-79	Sequence 79, Appl
74	33	100.0	250	7	US-11-096-568A-18242	Sequence 18242, A
75	33	100.0	251	7	US-11-087-099-1426	Sequence 1426, Ap
76	33	100.0	252	7	US-11-096-568A-20380	Sequence 20380, A
77	33	100.0	253	7	US-11-096-568A-31126	Sequence 31126, A
78	33	100.0	256	7	US-11-116-881A-492	Sequence 492, App
79	33	100.0	260	7	US-11-096-568A-24256	Sequence 24256, A
80	33	100.0	270	6	US-10-453-372-462	Sequence 462, App
81	33	100.0	270	6	US-10-453-372-464	Sequence 464, App
82	33	100.0	271	7	US-11-096-568A-15685	Sequence 15685, A
83	33	100.0	272	7	US-11-037-243-99	Sequence 99, Appl
84	33	100.0	272	7	US-11-096-568A-18241	Sequence 18241, A
85	33	100.0	274	7	US-11-072-512-2191	Sequence 2191, Ap
86	33	100.0	274	7	US-11-096-568A-15773	Sequence 15773, A
87	33	100.0	274	7	US-11-096-568A-20855	Sequence 20855, A
88	33	100.0	275	7	US-11-087-099-6038	Sequence 6038, Ap
89	33	100.0	276	7	US-11-096-568A-15772	Sequence 15772, A
90	33	100.0	280	7	US-11-096-568A-15120	Sequence 15120, A
91	33	100.0	281	6	US-10-934-944-387	Sequence 387, App
92	33	100.0	281	6	US-10-980-556-12	Sequence 12, Appl
93	33	100.0	281	7	US-11-116-881A-2298	Sequence 2298, Ap
94	33	100.0	282	6	US-10-980-556-16	Sequence 16, Appl
95	33	100.0	283	7	US-11-096-568A-28790	Sequence 28790, A
96	33	100.0	286	7	US-11-087-099-5131	Sequence 5131, Ap
97	33	100.0	286	7	US-11-096-568A-15684	Sequence 15684, A
98	33	100.0	293	7	US-11-096-568A-29769	Sequence 29769, A

99	33	100.0	299	7	US-11-096-568A-20853	Sequence 20853, A
100	33	100.0	312	7	US-11-096-568A-34393	Sequence 34393, A

ALIGNMENTS

```
RESULT 1
US-11-166-609-24
; Sequence 24, Application US/11166609
; Publication No. US20060015968A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIM
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  TRIMNELL, MARY
;  TITLE OF INVENTION: NUCLEOTIDE SEQUENCES MEDIATING MALE FERTILITY AND
;  TITLE OF INVENTION:  METHOD OF USING SAME
;  FILE REFERENCE: 1148CR
;  CURRENT APPLICATION NUMBER: US/11/166,609
;  CURRENT FILING DATE:  2005-06-24
;  PRIOR APPLICATION NUMBER: 10/412,000
;  PRIOR FILING DATE: 2003-04-11
;  PRIOR APPLICATION NUMBER: 09/670,153
;  PRIOR FILING DATE: 2000-09-26
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 24
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (2)..(3)
;   OTHER INFORMATION: Variable amino acid
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (5)
;   OTHER INFORMATION: Variable amino acid
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (7)
;   OTHER INFORMATION: Variable amino acid
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (9)
;   OTHER INFORMATION: Variable amino acid
US-11-166-609-24

  Query Match             100.0%;  Score 33;  DB 7;  Length 10;
  Best Local Similarity   90.0%;  Pred. No. 19;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |||||:||||
Db          1 FXXGXRXCXG 10


RESULT 2
US-11-166-609-25
; Sequence 25, Application US/11166609
; Publication No. US20060015968A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIM
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  TRIMNELL, MARY
;  TITLE OF INVENTION: NUCLEOTIDE SEQUENCES MEDIATING MALE FERTILITY AND
;  TITLE OF INVENTION:  METHOD OF USING SAME
;  FILE REFERENCE: 1148CR
;  CURRENT APPLICATION NUMBER: US/11/166,609
;  CURRENT FILING DATE:  2005-06-24
;  PRIOR APPLICATION NUMBER: 10/412,000
;  PRIOR FILING DATE: 2003-04-11
;  PRIOR APPLICATION NUMBER: 09/670,153
;  PRIOR FILING DATE: 2000-09-26
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 25
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Zea mays
US-11-166-609-25

  Query Match             100.0%;  Score 33;  DB 7;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 19;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FQAGPRICLG 10


RESULT 3
US-11-205-109-43
; Sequence 43, Application US/11205109
; Publication No. US20050287641A1
; GENERAL INFORMATION:
;  APPLICANT: Farnet, Chris
;  APPLICANT:  Zazopoulos, Emmanuel
;  APPLICANT:  Staffa, Alfredo
;  TITLE OF INVENTION: GENE CLUSTER FOR RAMOPLANIN BIOSYNTHESIS
;  FILE REFERENCE: 3002-2US
;  CURRENT APPLICATION NUMBER: US/11/205,109
;  CURRENT FILING DATE:  2005-08-17
;  PRIOR APPLICATION NUMBER: US/09/976,059
;  PRIOR FILING DATE: 2001-10-15
;  PRIOR APPLICATION NUMBER: US 60/239,924
;  PRIOR FILING DATE: 2000-10-13
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 43
;   LENGTH: 52
;   TYPE: PRT
;   ORGANISM: bacteria
US-11-205-109-43

  Query Match             100.0%;  Score 33;  DB 7;  Length 52;
  Best Local Similarity   40.0%;  Pred. No. 76;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         21 FGAGPRNCLG 30


RESULT 4
US-11-096-568A-25378
; Sequence 25378, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 25378
;   LENGTH: 61
;   TYPE: PRT
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(61)
;   OTHER INFORMATION: Ceres Seq. ID no. 12589527
```

US-11-096-568A-25378

Query Match 100.0%; Score 33; DB 7; Length 61;
Best Local Similarity 40.0%; Pred. No. 87;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          3 FGGGLRFCVG 12
```

RESULT 5
US-11-087-099-10778
; Sequence 10778, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 10778
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(71)
; OTHER INFORMATION: unsure at all Xaa locations
US-11-087-099-10778

Query Match 100.0%; Score 33; DB 7; Length 71;
Best Local Similarity 40.0%; Pred. No. 99;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         23 FKVGELPCRG 32
```

RESULT 6
US-11-116-881A-542
; Sequence 542, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Dongmei
; APPLICANT: Nielsen, Mark T.
; TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
; FILE REFERENCE: 07678/141014
; CURRENT APPLICATION NUMBER: US/11/116,881A
; CURRENT FILING DATE: 2005-04-27
; PRIOR APPLICATION NUMBER: 60/665,451
; PRIOR FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: 60/665,097
; PRIOR FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: 60/646,764
; PRIOR FILING DATE: 2005-01-25
; PRIOR APPLICATION NUMBER: 60/607,357
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 60/566,235
; PRIOR FILING DATE: 2004-04-29
; PRIOR APPLICATION NUMBER: 10/934,944
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 10/943,507
; PRIOR FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: 60/503,989
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/485,368
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 60/418,933
; PRIOR FILING DATE: 2002-10-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2300
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 542
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Nicotiana tabacum
US-11-116-881A-542

Query Match 100.0%; Score 33; DB 7; Length 73;
Best Local Similarity 40.0%; Pred. No. 1e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          6 FGSGRRSCPG 15
```

RESULT 7
US-11-116-881A-546
; Sequence 546, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Dongmei
; APPLICANT: Nielsen, Mark T.
; TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
; FILE REFERENCE: 07678/141014
; CURRENT APPLICATION NUMBER: US/11/116,881A
; CURRENT FILING DATE: 2005-04-27
; PRIOR APPLICATION NUMBER: 60/665,451
; PRIOR FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: 60/665,097
; PRIOR FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: 60/646,764
; PRIOR FILING DATE: 2005-01-25
; PRIOR APPLICATION NUMBER: 60/607,357
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 60/566,235
; PRIOR FILING DATE: 2004-04-29
; PRIOR APPLICATION NUMBER: 10/934,944
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 10/943,507
; PRIOR FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: 60/503,989
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/485,368
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 60/418,933
; PRIOR FILING DATE: 2002-10-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2300
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 546
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Nicotiana tabacum
US-11-116-881A-546

Query Match 100.0%; Score 33; DB 7; Length 73;
Best Local Similarity 40.0%; Pred. No. 1e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          6 FGSGRRSCPG 15
```

RESULT 8
US-10-467-657-8472
; Sequence 8472, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita

```
;  APPLICANT: PIZZA Mariagrazia
;  APPLICANT: MASIGNANI Vega
;  APPLICANT: MONACI Elisabetta
;  TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
;  FILE REFERENCE:
;  CURRENT APPLICATION NUMBER: US/10/467,657
;  CURRENT FILING DATE:  2003-08-11
;  PRIOR APPLICATION NUMBER: GB-0103424.8
;  PRIOR FILING DATE: 2001-02-12
;  NUMBER OF SEQ ID NOS: 9218
;  SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 8472
;   LENGTH: 75
;   TYPE: PRT
;   ORGANISM: Neisseria gonorrhoeae
US-10-467-657-8472

  Query Match             100.0%;  Score 33;  DB 6;  Length 75;
  Best Local Similarity   40.0%;  Pred. No. 1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         53 FLFGLQTCRG 62


RESULT 9
US-10-453-372-494
; Sequence 494, Application US/10453372
; Publication No. US20060003323A1
; GENERAL INFORMATION:
;  APPLICANT: Alsobrook, et al.
;  TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
;  FILE REFERENCE: 21402-589 A
;  CURRENT APPLICATION NUMBER: US/10/453,372
;  CURRENT FILING DATE:  2003-06-03
;  PRIOR APPLICATION NUMBER: 09/789390
;  PRIOR FILING DATE: 2001-02-23
;  PRIOR APPLICATION NUMBER: 60/185967
;  PRIOR FILING DATE: 2000-03-01
;  PRIOR APPLICATION NUMBER: 09/823187
;  PRIOR FILING DATE: 2001-03-29
;  PRIOR APPLICATION NUMBER: 60/195792
;  PRIOR FILING DATE: 2000-03-10
;  PRIOR APPLICATION NUMBER: 09/839446
;  PRIOR FILING DATE: 2001-03-19
;  PRIOR APPLICATION NUMBER: 60/199476
;  PRIOR FILING DATE: 2000-03-25
;  PRIOR APPLICATION NUMBER: 09/863776
;  PRIOR FILING DATE: 2001-05-23
;  PRIOR APPLICATION NUMBER: 60/208263
;  PRIOR FILING DATE: 2000-05-31
;  PRIOR APPLICATION NUMBER: 09/939398
;  PRIOR FILING DATE: 2001-08-24
;  PRIOR APPLICATION NUMBER: 60/227800
;  PRIOR FILING DATE: 2000-08-25
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1609
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 494
;   LENGTH: 84
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-453-372-494

  Query Match             100.0%;  Score 33;  DB 6;  Length 84;
  Best Local Similarity   40.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         51 FPQGSGRCLG 60


RESULT 10
US-11-116-881A-548
; Sequence 548, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Dongmei
;  APPLICANT:  Nielsen, Mark  T.
;  TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
;  FILE REFERENCE: 07678/141014
;  CURRENT APPLICATION NUMBER: US/11/116,881A
;  CURRENT FILING DATE:  2005-04-27
;  PRIOR APPLICATION NUMBER: 60/665,451
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/665,097
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/646,764
;  PRIOR FILING DATE: 2005-01-25
;  PRIOR APPLICATION NUMBER: 60/607,357
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 60/566,235
;  PRIOR FILING DATE: 2004-04-29
;  PRIOR APPLICATION NUMBER: 10/934,944
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 10/943,507
;  PRIOR FILING DATE: 2004-09-17
;  PRIOR APPLICATION NUMBER: 60/503,989
;  PRIOR FILING DATE: 2003-09-18
;  PRIOR APPLICATION NUMBER: 60/485,368
;  PRIOR FILING DATE: 2003-07-08
;  PRIOR APPLICATION NUMBER: 60/418,933
;  PRIOR FILING DATE: 2002-10-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2300
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 548
;   LENGTH: 86
;   TYPE: PRT
;   ORGANISM: Nicotiana tabacum
US-11-116-881A-548

  Query Match             100.0%;  Score 33;  DB 7;  Length 86;
  Best Local Similarity   40.0%;  Pred. No. 1.2e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         19 FGSGRRSCPG 28


RESULT 11
US-11-096-568A-25377
; Sequence 25377, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides I
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 25377
;   LENGTH: 99
;   TYPE: PRT
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(99)
;   OTHER INFORMATION: Ceres Seq. ID no. 12589526
US-11-096-568A-25377

  Query Match             100.0%;  Score 33;  DB 7;  Length 99;
```

Best Local Similarity 40.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         41 FGGGLRFCVG 50
```

RESULT 12
US-11-116-881A-2299
; Sequence 2299, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Dongmei
; APPLICANT: Nielsen, Mark T.
; TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
; FILE REFERENCE: 07678/141014
; CURRENT APPLICATION NUMBER: US/11/116,881A
; CURRENT FILING DATE: 2005-04-27
; PRIOR APPLICATION NUMBER: 60/665,451
; PRIOR FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: 60/665,097
; PRIOR FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: 60/646,764
; PRIOR FILING DATE: 2005-01-25
; PRIOR APPLICATION NUMBER: 60/607,357
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 60/566,235
; PRIOR FILING DATE: 2004-04-29
; PRIOR APPLICATION NUMBER: 10/934,944
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 10/943,507
; PRIOR FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: 60/503,989
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/485,368
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 60/418,933
; PRIOR FILING DATE: 2002-10-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2300
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2299
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Nicotiana Tabacam
US-11-116-881A-2299

Query Match 100.0%; Score 33; DB 7; Length 101;
Best Local Similarity 40.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         34 FGSGRRSCPG 43
```

RESULT 13
US-11-072-512-3124
; Sequence 3124, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3124
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-3124

Query Match 100.0%; Score 33; DB 7; Length 103;
Best Local Similarity 40.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         44 FSAGPRNCIG 53
```

RESULT 14
US-11-096-568A-26958
; Sequence 26958, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides F
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 26958
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(107)
; OTHER INFORMATION: Ceres Seq. ID no. 13636171
US-11-096-568A-26958

Query Match 100.0%; Score 33; DB 7; Length 107;
Best Local Similarity 40.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         40 FRKGKEACSG 49
```

RESULT 15
US-10-821-234-1478
; Sequence 1478, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia

```
;  FILE REFERENCE: 821A
;  CURRENT APPLICATION NUMBER: US/10/821,234
;  CURRENT FILING DATE:  2004-04-07
;  PRIOR APPLICATION NUMBER: US 60/462,047
;  PRIOR FILING DATE: 2003-04-07
;  NUMBER OF SEQ ID NOS: 1704
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 1478
;   LENGTH: 113
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-821-234-1478

  Query Match            100.0%;  Score 33;  DB 6;  Length 113;
  Best Local Similarity   40.0%;  Pred. No. 1.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         56 FLSGFISCVG 65


RESULT 16
US-11-087-099-6613
; Sequence 6613, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
;  APPLICANT: Abad, Mark S. et al.
;  TITLE OF INVENTION: Genes and Uses for Plant Improvement
;  FILE REFERENCE: 38-21(53450)B EP
;  CURRENT APPLICATION NUMBER: US/11/087,099
;  CURRENT FILING DATE:  2005-03-22
;  NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 6613
;   LENGTH: 117
;   TYPE: PRT
;   ORGANISM: Triticum aestivum
US-11-087-099-6613

  Query Match            100.0%;  Score 33;  DB 7;  Length 117;
  Best Local Similarity   40.0%;  Pred. No. 1.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db        108 FDGGADSCRG 117


RESULT 17
US-11-116-881A-464
; Sequence 464, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Dongmei
;  APPLICANT:  Nielsen, Mark  T.
;  TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
;  FILE REFERENCE: 07678/141014
;  CURRENT APPLICATION NUMBER: US/11/116,881A
;  CURRENT FILING DATE:  2005-04-27
;  PRIOR APPLICATION NUMBER: 60/665,451
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/665,097
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/646,764
;  PRIOR FILING DATE: 2005-01-25
;  PRIOR APPLICATION NUMBER: 60/607,357
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 60/566,235
;  PRIOR FILING DATE: 2004-04-29
;  PRIOR APPLICATION NUMBER: 10/934,944
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 10/943,507
;  PRIOR FILING DATE: 2004-09-17
;  PRIOR APPLICATION NUMBER: 60/503,989
;  PRIOR FILING DATE: 2003-09-18
;  PRIOR APPLICATION NUMBER: 60/485,368
;  PRIOR FILING DATE: 2003-07-08
;  PRIOR APPLICATION NUMBER: 60/418,933
;  PRIOR FILING DATE: 2002-10-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2300
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 464
;   LENGTH: 118
;   TYPE: PRT
;   ORGANISM: Nicotiana tabacum
US-11-116-881A-464

  Query Match            100.0%;  Score 33;  DB 7;  Length 118;
  Best Local Similarity   40.0%;  Pred. No. 1.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         51 FGSGRRSCPG 60


RESULT 18
US-11-116-881A-460
; Sequence 460, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Dongmei
;  APPLICANT:  Nielsen, Mark  T.
;  TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
;  FILE REFERENCE: 07678/141014
;  CURRENT APPLICATION NUMBER: US/11/116,881A
;  CURRENT FILING DATE:  2005-04-27
;  PRIOR APPLICATION NUMBER: 60/665,451
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/665,097
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/646,764
;  PRIOR FILING DATE: 2005-01-25
;  PRIOR APPLICATION NUMBER: 60/607,357
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 60/566,235
;  PRIOR FILING DATE: 2004-04-29
;  PRIOR APPLICATION NUMBER: 10/934,944
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 10/943,507
;  PRIOR FILING DATE: 2004-09-17
;  PRIOR APPLICATION NUMBER: 60/503,989
;  PRIOR FILING DATE: 2003-09-18
;  PRIOR APPLICATION NUMBER: 60/485,368
;  PRIOR FILING DATE: 2003-07-08
;  PRIOR APPLICATION NUMBER: 60/418,933
;  PRIOR FILING DATE: 2002-10-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2300
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 460
;   LENGTH: 119
;   TYPE: PRT
;   ORGANISM: Nicotiana tabacum
US-11-116-881A-460

  Query Match            100.0%;  Score 33;  DB 7;  Length 119;
  Best Local Similarity   40.0%;  Pred. No. 1.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         52 FGSGRRSCPG 61
```

```
RESULT 19
US-11-072-512-2610
; Sequence 2610, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: 084335-0191
;  CURRENT APPLICATION NUMBER: US/11/072,512
;  CURRENT FILING DATE:  2005-03-07
;  PRIOR APPLICATION NUMBER: US 60/350,978
;  PRIOR FILING DATE: 2002-01-25
;  PRIOR APPLICATION NUMBER: JP 2001-379298
;  PRIOR FILING DATE: 2001-11-05
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2610
;   LENGTH: 126
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-072-512-2610

  Query Match             100.0%;  Score 33;  DB 7;  Length 126;
  Best Local Similarity   40.0%;  Pred. No. 1.6e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         37 FSHGAPGCWG 46


RESULT 20
US-11-166-609-14
; Sequence 14, Application US/11166609
; Publication No. US20060015968A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIM
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  TRIMNELL, MARY
;  TITLE OF INVENTION: NUCLEOTIDE SEQUENCES MEDIATING MALE FERTILITY AND
;  TITLE OF INVENTION:  METHOD OF USING SAME
;  FILE REFERENCE: 1148CR
;  CURRENT APPLICATION NUMBER: US/11/166,609
;  CURRENT FILING DATE:  2005-06-24
;  PRIOR APPLICATION NUMBER: 10/412,000
;  PRIOR FILING DATE: 2003-04-11
;  PRIOR APPLICATION NUMBER: 09/670,153
;  PRIOR FILING DATE: 2000-09-26
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 14
;   LENGTH: 128
;   TYPE: PRT
;   ORGANISM: Zea mays
US-11-166-609-14

  Query Match             100.0%;  Score 33;  DB 7;  Length 128;
  Best Local Similarity   40.0%;  Pred. No. 1.6e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         68 FQAGPRICLG 77


RESULT 21
US-11-166-609-15
; Sequence 15, Application US/11166609
; Publication No. US20060015968A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTSEN, MARC C.
;  APPLICANT:  FOX, TIM
;  APPLICANT:  HUFFMAN, GARY
;  APPLICANT:  TRIMNELL, MARY
;  TITLE OF INVENTION: NUCLEOTIDE SEQUENCES MEDIATING MALE FERTILITY AND
;  TITLE OF INVENTION:  METHOD OF USING SAME
;  FILE REFERENCE: 1148CR
;  CURRENT APPLICATION NUMBER: US/11/166,609
;  CURRENT FILING DATE:  2005-06-24
;  PRIOR APPLICATION NUMBER: 10/412,000
;  PRIOR FILING DATE: 2003-04-11
;  PRIOR APPLICATION NUMBER: 09/670,153
;  PRIOR FILING DATE: 2000-09-26
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 15
;   LENGTH: 128
;   TYPE: PRT
;   ORGANISM: Zea mays
US-11-166-609-15

  Query Match             100.0%;  Score 33;  DB 7;  Length 128;
  Best Local Similarity   40.0%;  Pred. No. 1.6e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         68 FQAGPRICLG 77


RESULT 22
US-11-116-881A-500
; Sequence 500, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Dongmei
;  APPLICANT:  Nielsen, Mark  T.
;  TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
;  FILE REFERENCE: 07678/141014
;  CURRENT APPLICATION NUMBER: US/11/116,881A
;  CURRENT FILING DATE:  2005-04-27
;  PRIOR APPLICATION NUMBER: 60/665,451
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/665,097
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/646,764
;  PRIOR FILING DATE: 2005-01-25
;  PRIOR APPLICATION NUMBER: 60/607,357
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 60/566,235
;  PRIOR FILING DATE: 2004-04-29
;  PRIOR APPLICATION NUMBER: 10/934,944
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 10/943,507
;  PRIOR FILING DATE: 2004-09-17
;  PRIOR APPLICATION NUMBER: 60/503,989
;  PRIOR FILING DATE: 2003-09-18
;  PRIOR APPLICATION NUMBER: 60/485,368
```

```
;  PRIOR FILING DATE: 2003-07-08
;  PRIOR APPLICATION NUMBER: 60/418,933
;  PRIOR FILING DATE: 2002-10-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2300
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 500
;   LENGTH: 129
;   TYPE: PRT
;   ORGANISM: Nicotiana tabacum
US-11-116-881A-500

  Query Match             100.0%;  Score 33;  DB 7;  Length 129;
  Best Local Similarity   40.0%;  Pred. No. 1.6e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         62 FGSGRRSCPG 71


RESULT 23
US-11-116-881A-516
; Sequence 516, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Dongmei
;  APPLICANT:  Nielsen, Mark  T.
;  TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
;  FILE REFERENCE: 07678/141014
;  CURRENT APPLICATION NUMBER: US/11/116,881A
;  CURRENT FILING DATE:  2005-04-27
;  PRIOR APPLICATION NUMBER: 60/665,451
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/665,097
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/646,764
;  PRIOR FILING DATE: 2005-01-25
;  PRIOR APPLICATION NUMBER: 60/607,357
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 60/566,235
;  PRIOR FILING DATE: 2004-04-29
;  PRIOR APPLICATION NUMBER: 10/934,944
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 10/943,507
;  PRIOR FILING DATE: 2004-09-17
;  PRIOR APPLICATION NUMBER: 60/503,989
;  PRIOR FILING DATE: 2003-09-18
;  PRIOR APPLICATION NUMBER: 60/485,368
;  PRIOR FILING DATE: 2003-07-08
;  PRIOR APPLICATION NUMBER: 60/418,933
;  PRIOR FILING DATE: 2002-10-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2300
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 516
;   LENGTH: 129
;   TYPE: PRT
;   ORGANISM: Nicotiana tabacum
US-11-116-881A-516

  Query Match             100.0%;  Score 33;  DB 7;  Length 129;
  Best Local Similarity   40.0%;  Pred. No. 1.6e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db        103 FGSGRRSCPG 112


RESULT 24
US-11-116-881A-540
; Sequence 540, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Dongmei
;  APPLICANT:  Nielsen, Mark  T.
;  TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
;  FILE REFERENCE: 07678/141014
;  CURRENT APPLICATION NUMBER: US/11/116,881A
;  CURRENT FILING DATE:  2005-04-27
;  PRIOR APPLICATION NUMBER: 60/665,451
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/665,097
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/646,764
;  PRIOR FILING DATE: 2005-01-25
;  PRIOR APPLICATION NUMBER: 60/607,357
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 60/566,235
;  PRIOR FILING DATE: 2004-04-29
;  PRIOR APPLICATION NUMBER: 10/934,944
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 10/943,507
;  PRIOR FILING DATE: 2004-09-17
;  PRIOR APPLICATION NUMBER: 60/503,989
;  PRIOR FILING DATE: 2003-09-18
;  PRIOR APPLICATION NUMBER: 60/485,368
;  PRIOR FILING DATE: 2003-07-08
;  PRIOR APPLICATION NUMBER: 60/418,933
;  PRIOR FILING DATE: 2002-10-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2300
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 540
;   LENGTH: 129
;   TYPE: PRT
;   ORGANISM: Nicotiana tabacum
US-11-116-881A-540

  Query Match             100.0%;  Score 33;  DB 7;  Length 129;
  Best Local Similarity   40.0%;  Pred. No. 1.6e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         62 FGSGRRSCPG 71


RESULT 25
US-11-143-980-68
; Sequence 68, Application US/11143980
; Publication No. US20050272133A1
; GENERAL INFORMATION:
;  APPLICANT: He, Min
;  APPLICANT:  Hucul, John
;  APPLICANT:  Haltli, Bradley A.
;  APPLICANT:  Wagenaar, Melissa M.
;  APPLICANT:  Graziani, Edmund
;  APPLICANT:  Summers, Mia
;  APPLICANT:  Kulowski, Kerry
;  APPLICANT:  Pong, Kevin
;  TITLE OF INVENTION: Biosynthetic Gene Cluster for the Production of a Complex
;  TITLE OF INVENTION:  Polyketide
;  FILE REFERENCE: AM-101426US
;  CURRENT APPLICATION NUMBER: US/11/143,980
;  CURRENT FILING DATE:  2005-06-03
;  PRIOR APPLICATION NUMBER: US 60/664,483
;  PRIOR FILING DATE: 2005-03-23
;  PRIOR APPLICATION NUMBER: US 60/576,895
;  PRIOR FILING DATE: 2004-06-03
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 68
;   LENGTH: 134
;   TYPE: PRT
```

```
;   ORGANISM: Streptomyces sp.
US-11-143-980-68

  Query Match             100.0%;  Score 33;  DB 7;  Length 134;
  Best Local Similarity   40.0%;  Pred. No. 1.7e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         58 FDSGLDVCCG 67


RESULT 26
US-11-116-881A-456
; Sequence 456, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Dongmei
;  APPLICANT:  Nielsen, Mark  T.
;  TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
;  FILE REFERENCE: 07678/141014
;  CURRENT APPLICATION NUMBER: US/11/116,881A
;  CURRENT FILING DATE:  2005-04-27
;  PRIOR APPLICATION NUMBER: 60/665,451
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/665,097
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/646,764
;  PRIOR FILING DATE: 2005-01-25
;  PRIOR APPLICATION NUMBER: 60/607,357
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 60/566,235
;  PRIOR FILING DATE: 2004-04-29
;  PRIOR APPLICATION NUMBER: 10/934,944
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 10/943,507
;  PRIOR FILING DATE: 2004-09-17
;  PRIOR APPLICATION NUMBER: 60/503,989
;  PRIOR FILING DATE: 2003-09-18
;  PRIOR APPLICATION NUMBER: 60/485,368
;  PRIOR FILING DATE: 2003-07-08
;  PRIOR APPLICATION NUMBER: 60/418,933
;  PRIOR FILING DATE: 2002-10-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2300
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 456
;   LENGTH: 135
;   TYPE: PRT
;   ORGANISM: Nicotiana tabacum
US-11-116-881A-456

  Query Match             100.0%;  Score 33;  DB 7;  Length 135;
  Best Local Similarity   40.0%;  Pred. No. 1.7e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         68 FGPGRRSCPG 77


RESULT 27
US-11-116-881A-520
; Sequence 520, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Dongmei
;  APPLICANT:  Nielsen, Mark  T.
;  TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
;  FILE REFERENCE: 07678/141014
;  CURRENT APPLICATION NUMBER: US/11/116,881A
;  CURRENT FILING DATE:  2005-04-27
;  PRIOR APPLICATION NUMBER: 60/665,451
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/665,097
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/646,764
;  PRIOR FILING DATE: 2005-01-25
;  PRIOR APPLICATION NUMBER: 60/607,357
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 60/566,235
;  PRIOR FILING DATE: 2004-04-29
;  PRIOR APPLICATION NUMBER: 10/934,944
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 10/943,507
;  PRIOR FILING DATE: 2004-09-17
;  PRIOR APPLICATION NUMBER: 60/503,989
;  PRIOR FILING DATE: 2003-09-18
;  PRIOR APPLICATION NUMBER: 60/485,368
;  PRIOR FILING DATE: 2003-07-08
;  PRIOR APPLICATION NUMBER: 60/418,933
;  PRIOR FILING DATE: 2002-10-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2300
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 520
;   LENGTH: 137
;   TYPE: PRT
;   ORGANISM: Nicotiana tabacum
US-11-116-881A-520

  Query Match             100.0%;  Score 33;  DB 7;  Length 137;
  Best Local Similarity   40.0%;  Pred. No. 1.7e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         71 FGVGRRICPG 80


RESULT 28
US-11-116-881A-462
; Sequence 462, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Dongmei
;  APPLICANT:  Nielsen, Mark  T.
;  TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
;  FILE REFERENCE: 07678/141014
;  CURRENT APPLICATION NUMBER: US/11/116,881A
;  CURRENT FILING DATE:  2005-04-27
;  PRIOR APPLICATION NUMBER: 60/665,451
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/665,097
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/646,764
;  PRIOR FILING DATE: 2005-01-25
;  PRIOR APPLICATION NUMBER: 60/607,357
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 60/566,235
;  PRIOR FILING DATE: 2004-04-29
;  PRIOR APPLICATION NUMBER: 10/934,944
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 10/943,507
;  PRIOR FILING DATE: 2004-09-17
;  PRIOR APPLICATION NUMBER: 60/503,989
;  PRIOR FILING DATE: 2003-09-18
;  PRIOR APPLICATION NUMBER: 60/485,368
;  PRIOR FILING DATE: 2003-07-08
;  PRIOR APPLICATION NUMBER: 60/418,933
;  PRIOR FILING DATE: 2002-10-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2300
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 462
;   LENGTH: 140
```

```
;   TYPE: PRT
;   ORGANISM: Nicotiana tabacum
US-11-116-881A-462

  Query Match           100.0%;  Score 33;  DB 7;  Length 140;
  Best Local Similarity   40.0%;  Pred. No. 1.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         73 FGSGRRSCPG 82


RESULT 29
US-11-084-508-22
; Sequence 22, Application US/11084508
; Publication No. US20050260737A1
; GENERAL INFORMATION:
;  APPLICANT: Rahman, Raja Noor Zaliha Abd.
;  APPLICANT:  Salleh, Abu Bakar
;  APPLICANT:  Basri, Mahiran
;  APPLICANT:  Hun, Chin John
;  TITLE OF INVENTION: Novel Lipase Gene from Bacillus sphaericus 205y
;  FILE REFERENCE: KAN-101
;  CURRENT APPLICATION NUMBER: US/11/084,508
;  CURRENT FILING DATE:  2005-03-18
;  PRIOR APPLICATION NUMBER: MY 20040958
;  PRIOR FILING DATE: 2004-03-18
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
;   LENGTH: 141
;   TYPE: PRT
;   ORGANISM: Moraxella sp.
US-11-084-508-22

  Query Match           100.0%;  Score 33;  DB 7;  Length 141;
  Best Local Similarity   40.0%;  Pred. No. 1.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         32 FFHGGGFCIG 41


RESULT 30
US-11-116-881A-448
; Sequence 448, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Dongmei
;  APPLICANT:  Nielsen, Mark  T.
;  TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
;  FILE REFERENCE: 07678/141014
;  CURRENT APPLICATION NUMBER: US/11/116,881A
;  CURRENT FILING DATE:  2005-04-27
;  PRIOR APPLICATION NUMBER: 60/665,451
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/665,097
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/646,764
;  PRIOR FILING DATE: 2005-01-25
;  PRIOR APPLICATION NUMBER: 60/607,357
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 60/566,235
;  PRIOR FILING DATE: 2004-04-29
;  PRIOR APPLICATION NUMBER: 10/934,944
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 10/943,507
;  PRIOR FILING DATE: 2004-09-17
;  PRIOR APPLICATION NUMBER: 60/503,989
;  PRIOR FILING DATE: 2003-09-18
;  PRIOR APPLICATION NUMBER: 60/485,368
;  PRIOR FILING DATE: 2003-07-08
;  PRIOR APPLICATION NUMBER: 60/418,933
;  PRIOR FILING DATE: 2002-10-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2300
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 448
;   LENGTH: 151
;   TYPE: PRT
;   ORGANISM: Nicotiana tabacum
US-11-116-881A-448

  Query Match           100.0%;  Score 33;  DB 7;  Length 151;
  Best Local Similarity   40.0%;  Pred. No. 1.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         84 FGSGRRSCPG 93


RESULT 31
US-11-116-881A-510
; Sequence 510, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Dongmei
;  APPLICANT:  Nielsen, Mark  T.
;  TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
;  FILE REFERENCE: 07678/141014
;  CURRENT APPLICATION NUMBER: US/11/116,881A
;  CURRENT FILING DATE:  2005-04-27
;  PRIOR APPLICATION NUMBER: 60/665,451
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/665,097
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/646,764
;  PRIOR FILING DATE: 2005-01-25
;  PRIOR APPLICATION NUMBER: 60/607,357
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 60/566,235
;  PRIOR FILING DATE: 2004-04-29
;  PRIOR APPLICATION NUMBER: 10/934,944
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 10/943,507
;  PRIOR FILING DATE: 2004-09-17
;  PRIOR APPLICATION NUMBER: 60/503,989
;  PRIOR FILING DATE: 2003-09-18
;  PRIOR APPLICATION NUMBER: 60/485,368
;  PRIOR FILING DATE: 2003-07-08
;  PRIOR APPLICATION NUMBER: 60/418,933
;  PRIOR FILING DATE: 2002-10-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2300
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 510
;   LENGTH: 151
;   TYPE: PRT
;   ORGANISM: Nicotiana tabacum
US-11-116-881A-510

  Query Match           100.0%;  Score 33;  DB 7;  Length 151;
  Best Local Similarity   40.0%;  Pred. No. 1.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         86 FGAGRRICPG 95


RESULT 32
US-11-096-568A-12468
; Sequence 12468, Application US/11096568A
```

```
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 12468
;   LENGTH: 155
;   TYPE: PRT
;   ORGANISM: Triticum aestivum
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(155)
;   OTHER INFORMATION: Ceres Seq. ID no. 14301815
US-11-096-568A-12468

  Query Match             100.0%;  Score 33;  DB 7;  Length 155;
  Best Local Similarity   40.0%;  Pred. No. 1.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          5 FPRGNRLCSG 14


RESULT 33
US-11-096-568A-14864
; Sequence 14864, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 14864
;   LENGTH: 156
;   TYPE: PRT
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(156)
;   OTHER INFORMATION: Ceres Seq. ID no. 12340683
US-11-096-568A-14864

  Query Match             100.0%;  Score 33;  DB 7;  Length 156;
  Best Local Similarity   40.0%;  Pred. No. 1.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         90 FSCGRHSCVG 99


RESULT 34
US-11-096-568A-14863
; Sequence 14863, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandrov, Nickolai et al.
;  TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
;  TITLE OF INVENTION:  Therby
;  FILE REFERENCE: 2750-1592PUS2
;  CURRENT APPLICATION NUMBER: US/11/096,568A
;  CURRENT FILING DATE:  2005-04-01
;  NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 14863
;   LENGTH: 158
;   TYPE: PRT
;   ORGANISM: Zea mays subsp. mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(158)
;   OTHER INFORMATION: Ceres Seq. ID no. 12340682
US-11-096-568A-14863

  Query Match             100.0%;  Score 33;  DB 7;  Length 158;
  Best Local Similarity   40.0%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         92 FSCGRHSCVG 101


RESULT 35
US-11-116-881A-486
; Sequence 486, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Dongmei
;  APPLICANT:  Nielsen, Mark  T.
;  TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
;  FILE REFERENCE: 07678/141014
;  CURRENT APPLICATION NUMBER: US/11/116,881A
;  CURRENT FILING DATE:  2005-04-27
;  PRIOR APPLICATION NUMBER: 60/665,451
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/665,097
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/646,764
;  PRIOR FILING DATE: 2005-01-25
;  PRIOR APPLICATION NUMBER: 60/607,357
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 60/566,235
;  PRIOR FILING DATE: 2004-04-29
;  PRIOR APPLICATION NUMBER: 10/934,944
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 10/943,507
;  PRIOR FILING DATE: 2004-09-17
;  PRIOR APPLICATION NUMBER: 60/503,989
;  PRIOR FILING DATE: 2003-09-18
;  PRIOR APPLICATION NUMBER: 60/485,368
;  PRIOR FILING DATE: 2003-07-08
;  PRIOR APPLICATION NUMBER: 60/418,933
;  PRIOR FILING DATE: 2002-10-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2300
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 486
;   LENGTH: 162
;   TYPE: PRT
;   ORGANISM: Nicotiana tabacum
US-11-116-881A-486

  Query Match             100.0%;  Score 33;  DB 7;  Length 162;
  Best Local Similarity   40.0%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         97 FGAGRRICPG 106


RESULT 36
US-11-116-881A-484
; Sequence 484, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Dongmei
;  APPLICANT:  Nielsen, Mark  T.
```

```
;  TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
;  FILE REFERENCE: 07678/141014
;  CURRENT APPLICATION NUMBER: US/11/116,881A
;  CURRENT FILING DATE:  2005-04-27
;  PRIOR APPLICATION NUMBER: 60/665,451
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/665,097
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/646,764
;  PRIOR FILING DATE: 2005-01-25
;  PRIOR APPLICATION NUMBER: 60/607,357
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 60/566,235
;  PRIOR FILING DATE: 2004-04-29
;  PRIOR APPLICATION NUMBER: 10/934,944
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 10/943,507
;  PRIOR FILING DATE: 2004-09-17
;  PRIOR APPLICATION NUMBER: 60/503,989
;  PRIOR FILING DATE: 2003-09-18
;  PRIOR APPLICATION NUMBER: 60/485,368
;  PRIOR FILING DATE: 2003-07-08
;  PRIOR APPLICATION NUMBER: 60/418,933
;  PRIOR FILING DATE: 2002-10-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2300
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 484
;    LENGTH: 164
;    TYPE: PRT
;    ORGANISM: Nicotiana tabacum
;    FEATURE:
;    NAME/KEY: VARIANT
;    LOCATION: (27)..(27)
;    OTHER INFORMATION: Xaa = Leu, Pro, His, or Arg.
US-11-116-881A-484

  Query Match             100.0%;  Score 33;  DB 7;  Length 164;
  Best Local Similarity   40.0%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         92 FGMGRRACPG 101


RESULT 37
US-11-116-881A-470
; Sequence 470, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Dongmei
;  APPLICANT:  Nielsen, Mark  T.
;  TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
;  FILE REFERENCE: 07678/141014
;  CURRENT APPLICATION NUMBER: US/11/116,881A
;  CURRENT FILING DATE:  2005-04-27
;  PRIOR APPLICATION NUMBER: 60/665,451
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/665,097
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/646,764
;  PRIOR FILING DATE: 2005-01-25
;  PRIOR APPLICATION NUMBER: 60/607,357
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 60/566,235
;  PRIOR FILING DATE: 2004-04-29
;  PRIOR APPLICATION NUMBER: 10/934,944
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 10/943,507
;  PRIOR FILING DATE: 2004-09-17
;  PRIOR APPLICATION NUMBER: 60/503,989
;  PRIOR FILING DATE: 2003-09-18
;  PRIOR APPLICATION NUMBER: 60/485,368
;  PRIOR FILING DATE: 2003-07-08
;  PRIOR APPLICATION NUMBER: 60/418,933
;  PRIOR FILING DATE: 2002-10-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2300
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 470
;    LENGTH: 175
;    TYPE: PRT
;    ORGANISM: Nicotiana tabacum
US-11-116-881A-470

  Query Match             100.0%;  Score 33;  DB 7;  Length 175;
  Best Local Similarity   40.0%;  Pred. No. 2.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db        108 FGSGRRSCPG 117


RESULT 38
US-11-120-308-106
; Sequence 106, Application US/11120308
; Publication No. US20060005277A1
; GENERAL INFORMATION:
;  APPLICANT: Famodu, Omolayo O.
;  APPLICANT:  Forge, Charlie
;  APPLICANT:  Miao, Guo-Hua
;  TITLE OF INVENTION: cDNAS Encoding Polypeptides
;  FILE REFERENCE: BB-1365 US NA
;  CURRENT APPLICATION NUMBER: US/11/120,308
;  CURRENT FILING DATE:  2005-05-02
;  PRIOR APPLICATION NUMBER: US/10/078,770
;  PRIOR FILING DATE: 2002-02-19
;  PRIOR APPLICATION NUMBER: 09/614,188
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: 60/143,400
;  PRIOR FILING DATE: 1999-07-12
;  PRIOR APPLICATION NUMBER: 60/153,534
;  PRIOR FILING DATE: 1999-09-13
;  PRIOR APPLICATION NUMBER: 60/161,223
;  PRIOR FILING DATE: 1999-10-22
;  PRIOR APPLICATION NUMBER: 60/159,878
;  PRIOR FILING DATE: 1999-10-15
;  PRIOR APPLICATION NUMBER: 60/157,401
;  PRIOR FILING DATE: 1999-10-01
;  PRIOR APPLICATION NUMBER: 60/143,419
;  PRIOR FILING DATE: 1999-07-12
;  PRIOR APPLICATION NUMBER: 60/143,409
;  PRIOR FILING DATE: 1999-07-12
;  NUMBER OF SEQ ID NOS: 196
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 106
;    LENGTH: 176
;    TYPE: PRT
;    ORGANISM: Glycine max
;    FEATURE:
;    NAME/KEY: UNSURE
;    LOCATION: (134)
;    FEATURE:
;    NAME/KEY: UNSURE
;    LOCATION: (141)
;    FEATURE:
;    NAME/KEY: UNSURE
;    LOCATION: (145)
;    FEATURE:
;    NAME/KEY: UNSURE
;    LOCATION: (149)
;    FEATURE:
;    NAME/KEY: UNSURE
;    LOCATION: (170)
```

```
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: (172)
US-11-120-308-106

  Query Match             100.0%;  Score 33;  DB 7;  Length 176;
  Best Local Similarity   40.0%;  Pred. No. 2.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         60 FINGGPPCQG 69


RESULT 39
US-11-116-881A-522
; Sequence 522, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Dongmei
;  APPLICANT:  Nielsen, Mark  T.
;  TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
;  FILE REFERENCE: 07678/141014
;  CURRENT APPLICATION NUMBER: US/11/116,881A
;  CURRENT FILING DATE:  2005-04-27
;  PRIOR APPLICATION NUMBER: 60/665,451
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/665,097
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/646,764
;  PRIOR FILING DATE: 2005-01-25
;  PRIOR APPLICATION NUMBER: 60/607,357
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 60/566,235
;  PRIOR FILING DATE: 2004-04-29
;  PRIOR APPLICATION NUMBER: 10/934,944
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 10/943,507
;  PRIOR FILING DATE: 2004-09-17
;  PRIOR APPLICATION NUMBER: 60/503,989
;  PRIOR FILING DATE: 2003-09-18
;  PRIOR APPLICATION NUMBER: 60/485,368
;  PRIOR FILING DATE: 2003-07-08
;  PRIOR APPLICATION NUMBER: 60/418,933
;  PRIOR FILING DATE: 2002-10-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2300
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 522
;   LENGTH: 176
;   TYPE: PRT
;   ORGANISM: Nicotiana tabacum
US-11-116-881A-522

  Query Match             100.0%;  Score 33;  DB 7;  Length 176;
  Best Local Similarity   40.0%;  Pred. No. 2.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db        107 FGSGRRSCPG 116


RESULT 40
US-11-116-881A-502
; Sequence 502, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
;  APPLICANT: Xu, Dongmei
;  APPLICANT:  Nielsen, Mark  T.
;  TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
;  FILE REFERENCE: 07678/141014
;  CURRENT APPLICATION NUMBER: US/11/116,881A
;  CURRENT FILING DATE:  2005-04-27
;  PRIOR APPLICATION NUMBER: 60/665,451
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/665,097
;  PRIOR FILING DATE: 2005-03-24
;  PRIOR APPLICATION NUMBER: 60/646,764
;  PRIOR FILING DATE: 2005-01-25
;  PRIOR APPLICATION NUMBER: 60/607,357
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 60/566,235
;  PRIOR FILING DATE: 2004-04-29
;  PRIOR APPLICATION NUMBER: 10/934,944
;  PRIOR FILING DATE: 2004-09-03
;  PRIOR APPLICATION NUMBER: 10/943,507
;  PRIOR FILING DATE: 2004-09-17
;  PRIOR APPLICATION NUMBER: 60/503,989
;  PRIOR FILING DATE: 2003-09-18
;  PRIOR APPLICATION NUMBER: 60/485,368
;  PRIOR FILING DATE: 2003-07-08
;  PRIOR APPLICATION NUMBER: 60/418,933
;  PRIOR FILING DATE: 2002-10-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 2300
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 502
;   LENGTH: 177
;   TYPE: PRT
;   ORGANISM: Nicotiana tabacum
US-11-116-881A-502

  Query Match             100.0%;  Score 33;  DB 7;  Length 177;
  Best Local Similarity   40.0%;  Pred. No. 2.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db        127 FGAGRRICPG 136


RESULT 41
US-10-467-657-1228
; Sequence 1228, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
;  APPLICANT: CHIRON SpA
;  APPLICANT: FONTANA Maria Rita
;  APPLICANT: PIZZA Mariagrazia
;  APPLICANT: MASIGNANI Vega
;  APPLICANT: MONACI Elisabetta
;  TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
;  FILE REFERENCE:
;  CURRENT APPLICATION NUMBER: US/10/467,657
;  CURRENT FILING DATE:  2003-08-11
;  PRIOR APPLICATION NUMBER: GB-0103424.8
;  PRIOR FILING DATE: 2001-02-12
;  NUMBER OF SEQ ID NOS: 9218
;  SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1228
;   LENGTH: 185
;   TYPE: PRT
;   ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1228

  Query Match             100.0%;  Score 33;  DB 6;  Length 185;
  Best Local Similarity   40.0%;  Pred. No. 2.2e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         24 FCCGRFGCFG 33


RESULT 42
```

US-11-116-881A-458
; Sequence 458, Application US/11116881A
; Publication No. US20060041949A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Dongmei
; APPLICANT: Nielsen, Mark T.
; TITLE OF INVENTION: Nicotiana Nucleic Acid Molecules and Uses Thereof
; FILE REFERENCE: 07678/141014
; CURRENT APPLICATION NUMBER: US/11/116,881A
; CURRENT FILING DATE: 2005-04-27
; PRIOR APPLICATION NUMBER: 60/665,451
; PRIOR FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: 60/665,097
; PRIOR FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: 60/646,764
; PRIOR FILING DATE: 2005-01-25
; PRIOR APPLICATION NUMBER: 60/607,357
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 60/566,235
; PRIOR FILING DATE: 2004-04-29
; PRIOR APPLICATION NUMBER: 10/934,944
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 10/943,507
; PRIOR FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: 60/503,989
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/485,368
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 60/418,933
; PRIOR FILING DATE: 2002-10-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2300
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 458
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Nicotiana tabacum
US-11-116-881A-458

Query Match 100.0%; Score 33; DB 7; Length 187;
Best Local Similarity 40.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|::|:::|:|
Db 120 FGSGRRSCPG 129

RESULT 43
US-11-096-568A-23881
; Sequence 23881, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 23881
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(194)
; OTHER INFORMATION: Ceres Seq. ID no. 12416120
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (131)..(131)
; OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-23881

Query Match 100.0%; Score 33; DB 7; Length 194;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|::|::||:|
Db 125 FGAGRRXCAG 134

RESULT 44
US-11-096-568A-18243
; Sequence 18243, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides I
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 18243
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(197)
; OTHER INFORMATION: Ceres Seq. ID no. 12363892
US-11-096-568A-18243

Query Match 100.0%; Score 33; DB 7; Length 197;
Best Local Similarity 40.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|::|:::|:|
Db 154 FGAGRRMCPG 163

RESULT 45
US-09-995-493-12
; Sequence 12, Application US/09995493
; Publication No. US20060035293A1
; GENERAL INFORMATION:
; APPLICANT: Handfield, Martin
; APPLICANT: Hillman, Jeffrey
; APPLICANT: Progulske-Fox, Ann
; TITLE OF INVENTION: Identification of Actinobacillus actinomycetemcomitans Antigens f
; TITLE OF INVENTION: in the Diagnosis, Treatment, and Monitoring of Periodontal Disea
; FILE REFERENCE: MBHB01-662
; CURRENT APPLICATION NUMBER: US/09/995,493
; CURRENT FILING DATE: 2001-11-28
; NUMBER OF SEQ ID NOS: 234
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Actinobacillus actinomycetemcomitans
US-09-995-493-12

Query Match 100.0%; Score 33; DB 5; Length 199;
Best Local Similarity 40.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|::|:::|:|
Db 102 FESGAGICYG 111

RESULT 46
US-11-087-099-7405

; Sequence 7405, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 7405
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Populus x canescens
US-11-087-099-7405

Query Match 100.0%; Score 33; DB 7; Length 208;
Best Local Similarity 40.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|::|:::|:|
Db 136 FGAGRRMCPG 145

RESULT 47
US-11-087-099-11148
; Sequence 11148, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 11148
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Phanerochaete chrysosporium
US-11-087-099-11148

Query Match 100.0%; Score 33; DB 7; Length 208;
Best Local Similarity 40.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|::|:::|:|
Db 143 FNAGPRICLG 152

RESULT 48
US-11-096-568A-23880
; Sequence 23880, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 23880
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(210)
; OTHER INFORMATION: Ceres Seq. ID no. 12416119
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (147)..(147)
; OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-23880

Query Match 100.0%; Score 33; DB 7; Length 210;
Best Local Similarity 50.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|::|::||:|
Db 141 FGAGRRXCAG 150

RESULT 49
US-11-087-099-2635
; Sequence 2635, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 2635
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Candida tropicalis
US-11-087-099-2635

Query Match 100.0%; Score 33; DB 7; Length 213;
Best Local Similarity 40.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|::|:::|:|
Db 153 FNGGPRICLG 162

RESULT 50
US-11-087-099-11719
; Sequence 11719, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 11719
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Candida maltosa
US-11-087-099-11719

Query Match 100.0%; Score 33; DB 7; Length 213;
Best Local Similarity 40.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|::|:::|:|
Db 153 FNGGPRICLG 162

Search completed: March 8, 2006, 11:28:48
Job time : 24 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2006, 11:24:14 ; Search time 48 Seconds
(without alignments)
17.224 Million cell updates/sec

Title: US-10-751-235-14
Perfect score: 33
Sequence: 1 FXXGXXXCXG 10

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	100.0	10	1	US-07-912-900-1	Sequence 1, Appli
2	33	100.0	10	1	US-08-285-309-1	Sequence 1, Appli
3	33	100.0	10	1	US-08-502-046-1	Sequence 1, Appli
4	33	100.0	10	2	US-08-948-564-21	Sequence 21, Appl
5	33	100.0	10	2	US-09-904-615-129	Sequence 129, App
6	33	100.0	10	2	US-10-054-988-129	Sequence 129, App
7	33	100.0	11	1	US-07-912-900-13	Sequence 13, Appl
8	33	100.0	11	1	US-07-912-900-14	Sequence 14, Appl
9	33	100.0	11	1	US-07-912-900-15	Sequence 15, Appl
10	33	100.0	11	1	US-08-285-309-13	Sequence 13, Appl
11	33	100.0	11	1	US-08-285-309-14	Sequence 14, Appl
12	33	100.0	11	1	US-08-285-309-15	Sequence 15, Appl
13	33	100.0	11	1	US-08-313-075A-12	Sequence 12, Appl
14	33	100.0	11	1	US-08-313-075A-13	Sequence 13, Appl
15	33	100.0	11	1	US-08-313-075A-14	Sequence 14, Appl
16	33	100.0	11	1	US-08-502-046-13	Sequence 13, Appl
17	33	100.0	11	1	US-08-502-046-14	Sequence 14, Appl
18	33	100.0	11	1	US-08-502-046-15	Sequence 15, Appl
19	33	100.0	11	2	US-09-499-302A-3	Sequence 3, Appli
20	33	100.0	13	2	US-08-948-564-20	Sequence 20, Appl
21	33	100.0	14	1	US-08-845-161A-10	Sequence 10, Appl
22	33	100.0	14	1	US-08-845-161A-19	Sequence 19, Appl
23	33	100.0	14	1	US-08-845-161A-20	Sequence 20, Appl
24	33	100.0	14	1	US-08-845-161A-21	Sequence 21, Appl
25	33	100.0	14	1	US-08-845-161A-22	Sequence 22, Appl
26	33	100.0	14	1	US-08-845-161A-23	Sequence 23, Appl
27	33	100.0	14	1	US-08-845-161A-24	Sequence 24, Appl
28	33	100.0	14	1	US-08-845-161A-25	Sequence 25, Appl
29	33	100.0	14	2	US-09-270-751-10	Sequence 10, Appl
30	33	100.0	14	2	US-09-270-751-19	Sequence 19, Appl
31	33	100.0	14	2	US-09-270-751-20	Sequence 20, Appl
32	33	100.0	14	2	US-09-270-751-21	Sequence 21, Appl
33	33	100.0	14	2	US-09-270-751-22	Sequence 22, Appl
34	33	100.0	14	2	US-09-270-751-23	Sequence 23, Appl
35	33	100.0	14	2	US-09-270-751-24	Sequence 24, Appl
36	33	100.0	14	2	US-09-270-751-25	Sequence 25, Appl
37	33	100.0	15	6	5204096-22	Patent No. 5204096
38	33	100.0	16	2	US-09-471-276-1257	Sequence 1257, Ap
39	33	100.0	19	1	US-08-102-863-2	Sequence 2, Appli
40	33	100.0	19	4	PCT-US92-10885-2	Sequence 2, Appli
41	33	100.0	20	1	US-08-484-135-27	Sequence 27, Appl
42	33	100.0	20	1	US-08-484-635-208	Sequence 208, App
43	33	100.0	20	1	US-08-484-631-208	Sequence 208, App
44	33	100.0	20	1	US-08-827-570-208	Sequence 208, App
45	33	100.0	20	2	US-08-724-466B-6	Sequence 6, Appli
46	33	100.0	20	2	US-08-724-466B-7	Sequence 7, Appli
47	33	100.0	20	2	US-08-724-466B-8	Sequence 8, Appli
48	33	100.0	20	2	US-08-724-466B-9	Sequence 9, Appli
49	33	100.0	20	2	US-08-724-466B-10	Sequence 10, Appl
50	33	100.0	20	2	US-08-882-164D-6	Sequence 6, Appli
51	33	100.0	20	2	US-08-882-164D-7	Sequence 7, Appli
52	33	100.0	20	2	US-08-882-164D-8	Sequence 8, Appli
53	33	100.0	20	2	US-08-882-164D-9	Sequence 9, Appli
54	33	100.0	20	2	US-08-882-164D-10	Sequence 10, Appl
55	33	100.0	20	2	US-09-449-064A-33	Sequence 33, Appl
56	33	100.0	20	2	US-09-668-482-6	Sequence 6, Appli
57	33	100.0	20	2	US-09-668-482-7	Sequence 7, Appli
58	33	100.0	20	2	US-09-668-482-8	Sequence 8, Appli
59	33	100.0	20	2	US-09-668-482-9	Sequence 9, Appli
60	33	100.0	20	2	US-09-668-482-10	Sequence 10, Appl
61	33	100.0	30	2	US-09-749-637A-126	Sequence 126, App
62	33	100.0	30	2	US-09-749-637A-129	Sequence 129, App
63	33	100.0	30	2	US-09-749-637A-283	Sequence 283, App
64	33	100.0	35	2	US-09-082-279B-487	Sequence 487, App
65	33	100.0	35	2	US-09-082-279B-488	Sequence 488, App
66	33	100.0	35	2	US-09-082-279B-489	Sequence 489, App
67	33	100.0	35	2	US-09-082-279B-490	Sequence 490, App
68	33	100.0	35	2	US-09-082-279B-491	Sequence 491, App
69	33	100.0	35	2	US-09-082-279B-492	Sequence 492, App
70	33	100.0	35	2	US-09-082-279B-493	Sequence 493, App
71	33	100.0	35	2	US-09-082-279B-523	Sequence 523, App
72	33	100.0	35	2	US-09-082-279B-524	Sequence 524, App
73	33	100.0	35	2	US-09-082-279B-525	Sequence 525, App
74	33	100.0	35	2	US-09-082-279B-526	Sequence 526, App
75	33	100.0	35	2	US-09-082-279B-527	Sequence 527, App
76	33	100.0	35	2	US-09-082-279B-528	Sequence 528, App
77	33	100.0	35	2	US-09-082-279B-529	Sequence 529, App
78	33	100.0	35	2	US-08-474-349A-489	Sequence 489, App
79	33	100.0	35	2	US-08-474-349A-490	Sequence 490, App
80	33	100.0	35	2	US-08-474-349A-491	Sequence 491, App
81	33	100.0	35	2	US-08-474-349A-492	Sequence 492, App
82	33	100.0	35	2	US-08-474-349A-493	Sequence 493, App
83	33	100.0	35	2	US-08-474-349A-494	Sequence 494, App
84	33	100.0	35	2	US-09-315-304B-487	Sequence 487, App
85	33	100.0	35	2	US-09-315-304B-488	Sequence 488, App
86	33	100.0	35	2	US-09-315-304B-489	Sequence 489, App
87	33	100.0	35	2	US-09-315-304B-490	Sequence 490, App
88	33	100.0	35	2	US-09-315-304B-491	Sequence 491, App
89	33	100.0	35	2	US-09-315-304B-492	Sequence 492, App
90	33	100.0	35	2	US-09-315-304B-493	Sequence 493, App
91	33	100.0	35	2	US-09-315-304B-523	Sequence 523, App
92	33	100.0	35	2	US-09-315-304B-524	Sequence 524, App
93	33	100.0	35	2	US-09-315-304B-525	Sequence 525, App
94	33	100.0	35	2	US-09-315-304B-526	Sequence 526, App
95	33	100.0	35	2	US-09-315-304B-527	Sequence 527, App
96	33	100.0	35	2	US-09-315-304B-528	Sequence 528, App
97	33	100.0	35	2	US-09-315-304B-529	Sequence 529, App
98	33	100.0	35	2	US-09-834-784-487	Sequence 487, App
99	33	100.0	35	2	US-09-834-784-488	Sequence 488, App
100	33	100.0	35	2	US-09-834-784-489	Sequence 489, App

ALIGNMENTS

```
RESULT 1
US-07-912-900-1
; Sequence 1, Application US/07912900
; Patent No. 5349125
;  GENERAL INFORMATION:
;    APPLICANT:  Holton, Timothy A.
;    APPLICANT:  Cornish, Edwina C.
;    APPLICANT:  Kovacic, Filippa
;    APPLICANT:  Tanaka, Yoshikazu
;    APPLICANT:  Lester, Diane R.
;    TITLE OF INVENTION:  GENETIC SEQUENCES ENCODING FLAVONOID
;    TITLE OF INVENTION:  PATHWAY ENZYMES AND USES THEREFOR
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/912,900
;      FILING DATE:  19920713
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  8633
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-07-912-900-1

  Query Match             100.0%;  Score 33;  DB 1;  Length 10;
  Best Local Similarity   90.0%;  Pred. No. 1e+02;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |||||:||||
Db          1 FXXGXRXCXG 10


RESULT 2
US-08-285-309-1
; Sequence 1, Application US/08285309
; Patent No. 5569832
;  GENERAL INFORMATION:
;    APPLICANT:  Holton, Timothy A.
;    APPLICANT:  Cornish, Edwina C.
;    APPLICANT:  Kovacic, Filippa
;    APPLICANT:  Tanaka, Yoshikazu
;    APPLICANT:  Lester, Diane R.
;    TITLE OF INVENTION:  GENETIC SEQUENCES ENCODING A 3,5'-
;    TITLE OF INVENTION:  HYDROXYLASE AND USES
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/285,309
;      FILING DATE:  03-AUG-1994
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  8633Z
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-285-309-1

  Query Match             100.0%;  Score 33;  DB 1;  Length 10;
  Best Local Similarity   90.0%;  Pred. No. 1e+02;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |||||:||||
Db          1 FXXGXRXCXG 10


RESULT 3
US-08-502-046-1
; Sequence 1, Application US/08502046
; Patent No. 5861487
;  GENERAL INFORMATION:
;    APPLICANT:  Holton, Timothy A.
;    APPLICANT:  Cornish, Edwina C.
;    APPLICANT:  Kovacic, Filippa
;    APPLICANT:  Tanaka, Yoshikazu
;    APPLICANT:  Lester, Diane R.
;    TITLE OF INVENTION:  GENETIC SEQUENCES ENCODING A 3,5'-
;    TITLE OF INVENTION:  HYDROXYLASE AND USES
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/502,046
;      FILING DATE:  14-JUL-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/285,309
```

```
;      FILING DATE:  03-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  8633Z
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-502-046-1

  Query Match           100.0%;  Score 33;  DB 1;  Length 10;
  Best Local Similarity   90.0%;  Pred. No. 1e+02;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |||||:||||
Db          1 FXXGXRXCXG 10


RESULT 4
US-08-948-564-21
; Sequence 21, Application US/08948564
; Patent No. 6121512
;  GENERAL INFORMATION:
;    APPLICANT:  Siminszky, Balazs
;    APPLICANT:  Dewey, Ralph E.
;    APPLICANT:  Corbin, Frederick T.
;    TITLE OF INVENTION:  No. 6121512el Cytochrome P-450 Constructs and
;    TITLE OF INVENTION:  Methods of Producing Herbicide-Resistant Transgenic Plants
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Virginia C. Bennett
;      STREET:  PO Box 37428
;      CITY:  Raleigh
;      STATE:  No. 6121512th Carolina
;      COUNTRY:  USA
;      ZIP:  27627
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/948,564
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bennett, Virginia C.
;      REGISTRATION NUMBER:  37,092
;      REFERENCE/DOCKET NUMBER:  5051-409
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-854-1400
;      TELEFAX:  919-854-1401
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-948-564-21

  Query Match           100.0%;  Score 33;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              ||||||||||
Db          1 FXXGXXXCXG 10


RESULT 5
US-09-904-615-129
; Sequence 129, Application US/09904615
; Patent No. 6566325
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 49 Human Secreted Proteins
;  FILE REFERENCE: PZ032P1
;  CURRENT APPLICATION NUMBER: US/09/904,615
;  CURRENT FILING DATE:  2001-07-16
;  PRIOR APPLICATION NUMBER: 09/511,554
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/097,917
;  PRIOR FILING DATE: 1998-08-25
;  PRIOR APPLICATION NUMBER: 60/098,634
;  PRIOR FILING DATE: 1998-08-31
;  NUMBER OF SEQ ID NOS: 170
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 129
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-904-615-129

  Query Match           100.0%;  Score 33;  DB 2;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FSLGRRHCLG 10


RESULT 6
US-10-054-988-129
; Sequence 129, Application US/10054988
; Patent No. 6953667
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 49 Human Secreted Proteins
;  FILE REFERENCE: PZ032P1
;  CURRENT APPLICATION NUMBER: US/10/054,988
;  CURRENT FILING DATE:  2002-01-25
;  PRIOR APPLICATION NUMBER: 09/904,615
;  PRIOR FILING DATE: 2001-07-16
;  PRIOR APPLICATION NUMBER: 09/511,554
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/097,917
;  PRIOR FILING DATE: 1998-08-25
;  PRIOR APPLICATION NUMBER: 60/098,634
;  PRIOR FILING DATE: 1998-08-31
;  NUMBER OF SEQ ID NOS: 170
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 129
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-054-988-129

  Query Match           100.0%;  Score 33;  DB 2;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FSLGRRHCLG 10
```

```
RESULT 7
US-07-912-900-13
; Sequence 13, Application US/07912900
; Patent No. 5349125
;  GENERAL INFORMATION:
;    APPLICANT:  Holton, Timothy A.
;    APPLICANT:  Cornish, Edwina C.
;    APPLICANT:  Kovacic, Filippa
;    APPLICANT:  Tanaka, Yoshikazu
;    APPLICANT:  Lester, Diane R.
;    TITLE OF INVENTION:  GENETIC SEQUENCES ENCODING FLAVONOID
;    TITLE OF INVENTION:  PATHWAY ENZYMES AND USES THEREFOR
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/912,900
;      FILING DATE:  19920713
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  8633
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-07-912-900-13

  Query Match             100.0%;  Score 33;  DB 1;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGAGRRGCPG 11


RESULT 8
US-07-912-900-14
; Sequence 14, Application US/07912900
; Patent No. 5349125
;  GENERAL INFORMATION:
;    APPLICANT:  Holton, Timothy A.
;    APPLICANT:  Cornish, Edwina C.
;    APPLICANT:  Kovacic, Filippa
;    APPLICANT:  Tanaka, Yoshikazu
;    APPLICANT:  Lester, Diane R.
;    TITLE OF INVENTION:  GENETIC SEQUENCES ENCODING FLAVONOID
;    TITLE OF INVENTION:  PATHWAY ENZYMES AND USES THEREFOR
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/912,900
;      FILING DATE:  19920713
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  8633
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-07-912-900-14

  Query Match             100.0%;  Score 33;  DB 1;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGAGKRICPG 11


RESULT 9
US-07-912-900-15
; Sequence 15, Application US/07912900
; Patent No. 5349125
;  GENERAL INFORMATION:
;    APPLICANT:  Holton, Timothy A.
;    APPLICANT:  Cornish, Edwina C.
;    APPLICANT:  Kovacic, Filippa
;    APPLICANT:  Tanaka, Yoshikazu
;    APPLICANT:  Lester, Diane R.
;    TITLE OF INVENTION:  GENETIC SEQUENCES ENCODING FLAVONOID
;    TITLE OF INVENTION:  PATHWAY ENZYMES AND USES THEREFOR
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/912,900
;      FILING DATE:  19920713
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  8633
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
```

```
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;   INFORMATION FOR SEQ ID NO:  15:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  11 amino acids
;       TYPE:  AMINO ACID
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-07-912-900-15

  Query Match             100.0%;  Score 33;  DB 1;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGSGRRICPG 11


RESULT 10
US-08-285-309-13
; Sequence 13, Application US/08285309
; Patent No. 5569832
;  GENERAL INFORMATION:
;    APPLICANT:  Holton, Timothy A.
;    APPLICANT:  Cornish, Edwina C.
;    APPLICANT:  Kovacic, Filippa
;    APPLICANT:  Tanaka, Yoshikazu
;    APPLICANT:  Lester, Diane R.
;    TITLE OF INVENTION:  GENETIC SEQUENCES ENCODING A 3,5'-
;    TITLE OF INVENTION:  HYDROXYLASE AND USES
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/285,309
;      FILING DATE:  03-AUG-1994
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  8633Z
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-285-309-13

  Query Match             100.0%;  Score 33;  DB 1;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGAGRRGCPG 11


RESULT 11
US-08-285-309-14
; Sequence 14, Application US/08285309
; Patent No. 5569832
;  GENERAL INFORMATION:
;    APPLICANT:  Holton, Timothy A.
;    APPLICANT:  Cornish, Edwina C.
;    APPLICANT:  Kovacic, Filippa
;    APPLICANT:  Tanaka, Yoshikazu
;    APPLICANT:  Lester, Diane R.
;    TITLE OF INVENTION:  GENETIC SEQUENCES ENCODING A 3,5'-
;    TITLE OF INVENTION:  HYDROXYLASE AND USES
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/285,309
;      FILING DATE:  03-AUG-1994
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  8633Z
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-285-309-14

  Query Match             100.0%;  Score 33;  DB 1;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGAGKRICPG 11


RESULT 12
US-08-285-309-15
; Sequence 15, Application US/08285309
; Patent No. 5569832
;  GENERAL INFORMATION:
;    APPLICANT:  Holton, Timothy A.
;    APPLICANT:  Cornish, Edwina C.
;    APPLICANT:  Kovacic, Filippa
;    APPLICANT:  Tanaka, Yoshikazu
;    APPLICANT:  Lester, Diane R.
;    TITLE OF INVENTION:  GENETIC SEQUENCES ENCODING A 3,5'-
;    TITLE OF INVENTION:  HYDROXYLASE AND USES
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
```

```
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/285,309
;      FILING DATE:  03-AUG-1994
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  8633Z
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-285-309-15

  Query Match             100.0%;  Score 33;  DB 1;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGSGRRICPG 11


RESULT 13
US-08-313-075A-12
; Sequence 12, Application US/08313075A
; Patent No. 5639870
;  GENERAL INFORMATION:
;    APPLICANT:  Holton, Timothy A.
;    APPLICANT:  Cornish, Edwina C.
;    APPLICANT:  Tanaka, Yoshikazu
;    TITLE OF INVENTION:  GENETIC SEQUENCES ENCODING FLAVONOID
;    TITLE OF INVENTION:  PATHWAY ENZYMES AND USES THEREFOR
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/313,075A
;      FILING DATE:  30-NOV-1994
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PL 1538/92
;      FILING DATE:  27-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PL 6698/93
;      FILING DATE:  07-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PCT/AU93/00127
;      FILING DATE:  25-MAR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9433
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-313-075A-12

  Query Match             100.0%;  Score 33;  DB 1;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGAGRRGCPG 11


RESULT 14
US-08-313-075A-13
; Sequence 13, Application US/08313075A
; Patent No. 5639870
;  GENERAL INFORMATION:
;    APPLICANT:  Holton, Timothy A.
;    APPLICANT:  Cornish, Edwina C.
;    APPLICANT:  Tanaka, Yoshikazu
;    TITLE OF INVENTION:  GENETIC SEQUENCES ENCODING FLAVONOID
;    TITLE OF INVENTION:  PATHWAY ENZYMES AND USES THEREFOR
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/313,075A
;      FILING DATE:  30-NOV-1994
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PL 1538/92
;      FILING DATE:  27-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PL 6698/93
;      FILING DATE:  07-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PCT/AU93/00127
;      FILING DATE:  25-MAR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9433
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  13:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-313-075A-13

  Query Match           100.0%;  Score 33;  DB 1;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGAGKRICPG 11


RESULT 15
US-08-313-075A-14
; Sequence 14, Application US/08313075A
; Patent No. 5639870
;  GENERAL INFORMATION:
;    APPLICANT:  Holton, Timothy A.
;    APPLICANT:  Cornish, Edwina C.
;    APPLICANT:  Tanaka, Yoshikazu
;    TITLE OF INVENTION:  GENETIC SEQUENCES ENCODING FLAVONOID
;    TITLE OF INVENTION:  PATHWAY ENZYMES AND USES THEREFOR
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/313,075A
;      FILING DATE:  30-NOV-1994
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PL 1538/92
;      FILING DATE:  27-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PL 6698/93
;      FILING DATE:  07-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PCT/AU93/00127
;      FILING DATE:  25-MAR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9433
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-313-075A-14

  Query Match           100.0%;  Score 33;  DB 1;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGSGRRICPG 11


RESULT 16
US-08-502-046-13
; Sequence 13, Application US/08502046
; Patent No. 5861487
;  GENERAL INFORMATION:
;    APPLICANT:  Holton, Timothy A.
;    APPLICANT:  Cornish, Edwina C.
;    APPLICANT:  Kovacic, Filippa
;    APPLICANT:  Tanaka, Yoshikazu
;    APPLICANT:  Lester, Diane R.
;    TITLE OF INVENTION:  GENETIC SEQUENCES ENCODING A 3,5'-
;    TITLE OF INVENTION:  HYDROXYLASE AND USES
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/502,046
;      FILING DATE:  14-JUL-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/285,309
;      FILING DATE:  03-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  8633Z
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-502-046-13

  Query Match           100.0%;  Score 33;  DB 1;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGAGRRGCPG 11


RESULT 17
US-08-502-046-14
; Sequence 14, Application US/08502046
; Patent No. 5861487
;  GENERAL INFORMATION:
;    APPLICANT:  Holton, Timothy A.
;    APPLICANT:  Cornish, Edwina C.
;    APPLICANT:  Kovacic, Filippa
;    APPLICANT:  Tanaka, Yoshikazu
```

```
;    APPLICANT:  Lester, Diane R.
;    TITLE OF INVENTION:  GENETIC SEQUENCES ENCODING A 3,5'-
;    TITLE OF INVENTION:  HYDROXYLASE AND USES
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/502,046
;      FILING DATE:  14-JUL-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/285,309
;      FILING DATE:  03-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  8633Z
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-502-046-14

  Query Match             100.0%;  Score 33;  DB 1;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGAGKRICPG 11


RESULT 18
US-08-502-046-15
; Sequence 15, Application US/08502046
; Patent No. 5861487
;  GENERAL INFORMATION:
;    APPLICANT:  Holton, Timothy A.
;    APPLICANT:  Cornish, Edwina C.
;    APPLICANT:  Kovacic, Filippa
;    APPLICANT:  Tanaka, Yoshikazu
;    APPLICANT:  Lester, Diane R.
;    TITLE OF INVENTION:  GENETIC SEQUENCES ENCODING A 3,5'-
;    TITLE OF INVENTION:  HYDROXYLASE AND USES
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/502,046
;      FILING DATE:  14-JUL-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/285,309
;      FILING DATE:  03-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  8633Z
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-502-046-15

  Query Match             100.0%;  Score 33;  DB 1;  Length 11;
  Best Local Similarity   40.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGSGRRICPG 11


RESULT 19
US-09-499-302A-3
; Sequence 3, Application US/09499302A
; Patent No. 6369212
; GENERAL INFORMATION:
;  APPLICANT: BOUNG-JUN, OH
;  APPLICANT:  MOON, KYUNG KO
;  APPLICANT:  YOUNG, SOON KIM
;  TITLE OF INVENTION: A CYTOCHROME P450 GENE HIGHLY EXPRESSED IN THE
;  TITLE OF INVENTION:  INCOMPATIBLE INTERACTION
;  FILE REFERENCE: 10324/P64443USO
;  CURRENT APPLICATION NUMBER: US/09/499,302A
;  CURRENT FILING DATE:  2000-02-07
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;    LENGTH: 11
;    TYPE: PRT
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence: Consensus
;    OTHER INFORMATION: heme-binding domain
;    NAME/KEY: MOD_RES
;    LOCATION: (4)
;    OTHER INFORMATION: variable or unknown amino acid
;    NAME/KEY: MOD_RES
;    LOCATION: (6)
;    OTHER INFORMATION: variable or unknown amino acid
;    NAME/KEY: MOD_RES
;    LOCATION: (8)
;    OTHER INFORMATION: variable or unknown amino acid
;    NAME/KEY: MOD_RES
;    LOCATION: (10)
;    OTHER INFORMATION: variable or unknown amino acid
US-09-499-302A-3

  Query Match             100.0%;  Score 33;  DB 2;  Length 11;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+02;
  Matches    8;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 FXXGXXXCXG 10
              |:|||:||||
Db          2 FGXGXRXCXG 11


RESULT 20
US-08-948-564-20
; Sequence 20, Application US/08948564
; Patent No. 6121512
;  GENERAL INFORMATION:
;    APPLICANT:  Siminszky, Balazs
;    APPLICANT:  Dewey, Ralph E.
;    APPLICANT:  Corbin, Frederick T.
;    TITLE OF INVENTION:  No. 6121512el Cytochrome P-450 Constructs and
;    TITLE OF INVENTION:  Methods of Producing Herbicide-Resistant Transgenic Plants
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Virginia C. Bennett
;      STREET:  PO Box 37428
;      CITY:  Raleigh
;      STATE:  No. 6121512th Carolina
;      COUNTRY:  USA
;      ZIP:  27627
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/948,564
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bennett, Virginia C.
;      REGISTRATION NUMBER:  37,092
;      REFERENCE/DOCKET NUMBER:  5051-409
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-854-1400
;      TELEFAX:  919-854-1401
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-948-564-20

  Query Match             100.0%;  Score 33;  DB 2;  Length 13;
  Best Local Similarity   80.0%;  Pred. No. 1.3e+02;
  Matches    8;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |:|||:||||
Db          4 FGXGXRXCXG 13


RESULT 21
US-08-845-161A-10
; Sequence 10, Application US/08845161A
; Patent No. 5976850
;  GENERAL INFORMATION:
;    APPLICANT:  Lathe, Richard
;    APPLICANT:  Rose, Kenneth A.
;    APPLICANT:  Stapleton, Genevieve
;    TITLE OF INVENTION:  HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;    TITLE OF INVENTION:  SEQUENCES CODING THEREFOR AND USED THEREOF
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON & VANDERHYE P.C.
;      STREET:  1100 No. 5976850th Glebe Rd.  8th floor
;      CITY:  Arlington
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22201-4741
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/845,161A
;      FILING DATE:  21-APR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB95/02465
;      FILING DATE:  18-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9421093.7
;      FILING DATE:  19-OCT-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wilson, Mary J.
;      REGISTRATION NUMBER:  32,955
;      REFERENCE/DOCKET NUMBER:  604-408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-816-4000
;      TELEFAX:  703-816-4100
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-845-161A-10

  Query Match             100.0%;  Score 33;  DB 1;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              ||||||||||
Db          1 FXXGXXXCXG 10


RESULT 22
US-08-845-161A-19
; Sequence 19, Application US/08845161A
; Patent No. 5976850
;  GENERAL INFORMATION:
;    APPLICANT:  Lathe, Richard
;    APPLICANT:  Rose, Kenneth A.
;    APPLICANT:  Stapleton, Genevieve
;    TITLE OF INVENTION:  HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;    TITLE OF INVENTION:  SEQUENCES CODING THEREFOR AND USED THEREOF
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON & VANDERHYE P.C.
;      STREET:  1100 No. 5976850th Glebe Rd.  8th floor
;      CITY:  Arlington
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22201-4741
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/845,161A
;      FILING DATE:  21-APR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB95/02465
;      FILING DATE:  18-OCT-1995
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9421093.7
;      FILING DATE:  19-OCT-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wilson, Mary J.
;      REGISTRATION NUMBER:  32,955
;      REFERENCE/DOCKET NUMBER:  604-408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-816-4000
;      TELEFAX:  703-816-4100
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-845-161A-19

  Query Match             100.0%;  Score 33;  DB 1;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGLGTSKCPG 10


RESULT 23
US-08-845-161A-20
; Sequence 20, Application US/08845161A
; Patent No. 5976850
;  GENERAL INFORMATION:
;    APPLICANT:  Lathe, Richard
;    APPLICANT:  Rose, Kenneth A.
;    APPLICANT:  Stapleton, Genevieve
;    TITLE OF INVENTION:  HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;    TITLE OF INVENTION:  SEQUENCES CODING THEREFOR AND USED THEREOF
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON & VANDERHYE P.C.
;      STREET:  1100 No. 5976850th Glebe Rd.  8th floor
;      CITY:  Arlington
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22201-4741
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/845,161A
;      FILING DATE:  21-APR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB95/02465
;      FILING DATE:  18-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9421093.7
;      FILING DATE:  19-OCT-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wilson, Mary J.
;      REGISTRATION NUMBER:  32,955
;      REFERENCE/DOCKET NUMBER:  604-408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-816-4000
;      TELEFAX:  703-816-4100
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-845-161A-20

  Query Match             100.0%;  Score 33;  DB 1;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGSGATICPG 10


RESULT 24
US-08-845-161A-21
; Sequence 21, Application US/08845161A
; Patent No. 5976850
;  GENERAL INFORMATION:
;    APPLICANT:  Lathe, Richard
;    APPLICANT:  Rose, Kenneth A.
;    APPLICANT:  Stapleton, Genevieve
;    TITLE OF INVENTION:  HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;    TITLE OF INVENTION:  SEQUENCES CODING THEREFOR AND USED THEREOF
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON & VANDERHYE P.C.
;      STREET:  1100 No. 5976850th Glebe Rd.  8th floor
;      CITY:  Arlington
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22201-4741
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/845,161A
;      FILING DATE:  21-APR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB95/02465
;      FILING DATE:  18-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9421093.7
;      FILING DATE:  19-OCT-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wilson, Mary J.
;      REGISTRATION NUMBER:  32,955
;      REFERENCE/DOCKET NUMBER:  604-408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-816-4000
;      TELEFAX:  703-816-4100
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-845-161A-21

  Query Match             100.0%;  Score 33;  DB 1;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGAGPRSCVG 10


RESULT 25
US-08-845-161A-22
```

```
; Sequence 22, Application US/08845161A
; Patent No. 5976850
;  GENERAL INFORMATION:
;    APPLICANT:  Lathe, Richard
;    APPLICANT:  Rose, Kenneth A.
;    APPLICANT:  Stapleton, Genevieve
;    TITLE OF INVENTION:  HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;    TITLE OF INVENTION:  SEQUENCES CODING THEREFOR AND USED THEREOF
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON & VANDERHYE P.C.
;      STREET:  1100 No. 5976850th Glebe Rd.  8th floor
;      CITY:  Arlington
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22201-4741
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/845,161A
;      FILING DATE:  21-APR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB95/02465
;      FILING DATE:  18-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9421093.7
;      FILING DATE:  19-OCT-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wilson, Mary J.
;      REGISTRATION NUMBER:  32,955
;      REFERENCE/DOCKET NUMBER:  604-408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-816-4000
;      TELEFAX:  703-816-4100
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-845-161A-22

  Query Match           100.0%;  Score 33;  DB 1;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGFGMRQCLG 10


RESULT 26
US-08-845-161A-23
; Sequence 23, Application US/08845161A
; Patent No. 5976850
;  GENERAL INFORMATION:
;    APPLICANT:  Lathe, Richard
;    APPLICANT:  Rose, Kenneth A.
;    APPLICANT:  Stapleton, Genevieve
;    TITLE OF INVENTION:  HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;    TITLE OF INVENTION:  SEQUENCES CODING THEREFOR AND USED THEREOF
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON & VANDERHYE P.C.
;      STREET:  1100 No. 5976850th Glebe Rd.  8th floor
;      CITY:  Arlington
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22201-4741
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/845,161A
;      FILING DATE:  21-APR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB95/02465
;      FILING DATE:  18-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9421093.7
;      FILING DATE:  19-OCT-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wilson, Mary J.
;      REGISTRATION NUMBER:  32,955
;      REFERENCE/DOCKET NUMBER:  604-408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-816-4000
;      TELEFAX:  703-816-4100
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-845-161A-23

  Query Match           100.0%;  Score 33;  DB 1;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGCGARVCLG 10


RESULT 27
US-08-845-161A-24
; Sequence 24, Application US/08845161A
; Patent No. 5976850
;  GENERAL INFORMATION:
;    APPLICANT:  Lathe, Richard
;    APPLICANT:  Rose, Kenneth A.
;    APPLICANT:  Stapleton, Genevieve
;    TITLE OF INVENTION:  HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;    TITLE OF INVENTION:  SEQUENCES CODING THEREFOR AND USED THEREOF
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON & VANDERHYE P.C.
;      STREET:  1100 No. 5976850th Glebe Rd.  8th floor
;      CITY:  Arlington
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22201-4741
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/845,161A
;      FILING DATE:  21-APR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB95/02465
;      FILING DATE:  18-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9421093.7
```

```
;      FILING DATE:  19-OCT-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wilson, Mary J.
;      REGISTRATION NUMBER:  32,955
;      REFERENCE/DOCKET NUMBER:  604-408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-816-4000
;      TELEFAX:  703-816-4100
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-845-161A-24

  Query Match             100.0%;  Score 33;  DB 1;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGWGVRQCLG 10


RESULT 28
US-08-845-161A-25
; Sequence 25, Application US/08845161A
; Patent No. 5976850
;  GENERAL INFORMATION:
;    APPLICANT:  Lathe, Richard
;    APPLICANT:  Rose, Kenneth A.
;    APPLICANT:  Stapleton, Genevieve
;    TITLE OF INVENTION:  HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;    TITLE OF INVENTION:  SEQUENCES CODING THEREFOR AND USED THEREOF
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIXON & VANDERHYE P.C.
;      STREET:  1100 No. 5976850th Glebe Rd.  8th floor
;      CITY:  Arlington
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22201-4741
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/845,161A
;      FILING DATE:  21-APR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB95/02465
;      FILING DATE:  18-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9421093.7
;      FILING DATE:  19-OCT-1994
;   ATTORNEY/AGENT INFORMATION:
;      NAME:  Wilson, Mary J.
;      REGISTRATION NUMBER:  32,955
;      REFERENCE/DOCKET NUMBER:  604-408
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-816-4000
;      TELEFAX:  703-816-4100
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-845-161A-25

  Query Match             100.0%;  Score 33;  DB 1;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGYGVRACLG 10


RESULT 29
US-09-270-751-10
; Sequence 10, Application US/09270751
; Patent No. 6184350
;  GENERAL INFORMATION:
;       APPLICANT: Lathe, Richard
;                  Rose, Kenneth A.
;                  Stapleton, Genevieve
;       TITLE OF INVENTION: HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;                           SEQUENCES CODING THEREFOR AND USED THEREOF
;       NUMBER OF SEQUENCES: 45
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: NIXON & VANDERHYE P.C.
;            STREET: 1100 No. 6184350th Glebe Rd.  8th floor
;            CITY: Arlington
;            STATE: VA
;            COUNTRY: USA
;            ZIP: 22201-4741
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/270,751
;            FILING DATE: 17-Apr-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: PCT/GB95/02465
;            FILING DATE: 18-OCT-1995
;            APPLICATION NUMBER: GB 9421093.7
;            FILING DATE: 19-OCT-1994
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Wilson, Mary J.
;            REGISTRATION NUMBER: 32,955
;            REFERENCE/DOCKET NUMBER: 604-408
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 703-816-4000
;            TELEFAX: 703-816-4100
;   INFORMATION FOR SEQ ID NO: 10:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 14 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: <Unknown>
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-270-751-10

  Query Match             100.0%;  Score 33;  DB 2;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+02;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              ||||||||||
Db          1 FXXGXXXCXG 10


RESULT 30
US-09-270-751-19
; Sequence 19, Application US/09270751
; Patent No. 6184350
```

```
;  GENERAL INFORMATION:
;       APPLICANT: Lathe, Richard
;                  Rose, Kenneth A.
;                  Stapleton, Genevieve
;       TITLE OF INVENTION: HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;                           SEQUENCES CODING THEREFOR AND USED THEREOF
;       NUMBER OF SEQUENCES: 45
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: NIXON & VANDERHYE P.C.
;            STREET: 1100 No. 6184350th Glebe Rd.  8th floor
;            CITY: Arlington
;            STATE: VA
;            COUNTRY: USA
;            ZIP: 22201-4741
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/270,751
;            FILING DATE: 17-Apr-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: PCT/GB95/02465
;            FILING DATE: 18-OCT-1995
;            APPLICATION NUMBER: GB 9421093.7
;            FILING DATE: 19-OCT-1994
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Wilson, Mary J.
;            REGISTRATION NUMBER: 32,955
;            REFERENCE/DOCKET NUMBER: 604-408
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 703-816-4000
;            TELEFAX: 703-816-4100
;  INFORMATION FOR SEQ ID NO: 19:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 14 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: <Unknown>
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-270-751-19

  Query Match           100.0%;  Score 33;  DB 2;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGLGTSKCPG 10


RESULT 31
US-09-270-751-20
; Sequence 20, Application US/09270751
; Patent No. 6184350
;  GENERAL INFORMATION:
;       APPLICANT: Lathe, Richard
;                  Rose, Kenneth A.
;                  Stapleton, Genevieve
;       TITLE OF INVENTION: HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;                           SEQUENCES CODING THEREFOR AND USED THEREOF
;       NUMBER OF SEQUENCES: 45
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: NIXON & VANDERHYE P.C.
;            STREET: 1100 No. 6184350th Glebe Rd.  8th floor
;            CITY: Arlington
;            STATE: VA
;            COUNTRY: USA
;            ZIP: 22201-4741
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/270,751
;            FILING DATE: 17-Apr-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: PCT/GB95/02465
;            FILING DATE: 18-OCT-1995
;            APPLICATION NUMBER: GB 9421093.7
;            FILING DATE: 19-OCT-1994
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Wilson, Mary J.
;            REGISTRATION NUMBER: 32,955
;            REFERENCE/DOCKET NUMBER: 604-408
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 703-816-4000
;            TELEFAX: 703-816-4100
;  INFORMATION FOR SEQ ID NO: 20:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 14 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: <Unknown>
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-270-751-20

  Query Match           100.0%;  Score 33;  DB 2;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGSGATICPG 10


RESULT 32
US-09-270-751-21
; Sequence 21, Application US/09270751
; Patent No. 6184350
;  GENERAL INFORMATION:
;       APPLICANT: Lathe, Richard
;                  Rose, Kenneth A.
;                  Stapleton, Genevieve
;       TITLE OF INVENTION: HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;                           SEQUENCES CODING THEREFOR AND USED THEREOF
;       NUMBER OF SEQUENCES: 45
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: NIXON & VANDERHYE P.C.
;            STREET: 1100 No. 6184350th Glebe Rd.  8th floor
;            CITY: Arlington
;            STATE: VA
;            COUNTRY: USA
;            ZIP: 22201-4741
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/270,751
;            FILING DATE: 17-Apr-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: PCT/GB95/02465
;            FILING DATE: 18-OCT-1995
;            APPLICATION NUMBER: GB 9421093.7
;            FILING DATE: 19-OCT-1994
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Wilson, Mary J.
```

```
;             REGISTRATION NUMBER: 32,955
;             REFERENCE/DOCKET NUMBER: 604-408
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-816-4000
;             TELEFAX: 703-816-4100
;   INFORMATION FOR SEQ ID NO: 21:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 14 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: <Unknown>
;             TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-270-751-21

  Query Match             100.0%;  Score 33;  DB 2;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGAGPRSCVG 10


RESULT 33
US-09-270-751-22
; Sequence 22, Application US/09270751
; Patent No. 6184350
;   GENERAL INFORMATION:
;       APPLICANT: Lathe, Richard
;                  Rose, Kenneth A.
;                  Stapleton, Genevieve
;       TITLE OF INVENTION: HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;                           SEQUENCES CODING THEREFOR AND USED THEREOF
;       NUMBER OF SEQUENCES: 45
;       CORRESPONDENCE ADDRESS:
;             ADDRESSEE: NIXON & VANDERHYE P.C.
;             STREET: 1100 No. 6184350th Glebe Rd.  8th floor
;             CITY: Arlington
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22201-4741
;       COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/270,751
;             FILING DATE: 17-Apr-2000
;             CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: PCT/GB95/02465
;             FILING DATE: 18-OCT-1995
;             APPLICATION NUMBER: GB 9421093.7
;             FILING DATE: 19-OCT-1994
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Wilson, Mary J.
;             REGISTRATION NUMBER: 32,955
;             REFERENCE/DOCKET NUMBER: 604-408
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-816-4000
;             TELEFAX: 703-816-4100
;   INFORMATION FOR SEQ ID NO: 22:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 14 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: <Unknown>
;             TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-270-751-22

  Query Match             100.0%;  Score 33;  DB 2;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGFGMRQCLG 10


RESULT 34
US-09-270-751-23
; Sequence 23, Application US/09270751
; Patent No. 6184350
;   GENERAL INFORMATION:
;       APPLICANT: Lathe, Richard
;                  Rose, Kenneth A.
;                  Stapleton, Genevieve
;       TITLE OF INVENTION: HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;                           SEQUENCES CODING THEREFOR AND USED THEREOF
;       NUMBER OF SEQUENCES: 45
;       CORRESPONDENCE ADDRESS:
;             ADDRESSEE: NIXON & VANDERHYE P.C.
;             STREET: 1100 No. 6184350th Glebe Rd.  8th floor
;             CITY: Arlington
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22201-4741
;       COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/270,751
;             FILING DATE: 17-Apr-2000
;             CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: PCT/GB95/02465
;             FILING DATE: 18-OCT-1995
;             APPLICATION NUMBER: GB 9421093.7
;             FILING DATE: 19-OCT-1994
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Wilson, Mary J.
;             REGISTRATION NUMBER: 32,955
;             REFERENCE/DOCKET NUMBER: 604-408
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-816-4000
;             TELEFAX: 703-816-4100
;   INFORMATION FOR SEQ ID NO: 23:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 14 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: <Unknown>
;             TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-270-751-23

  Query Match             100.0%;  Score 33;  DB 2;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGCGARVCLG 10


RESULT 35
US-09-270-751-24
; Sequence 24, Application US/09270751
; Patent No. 6184350
;   GENERAL INFORMATION:
;       APPLICANT: Lathe, Richard
```

```
;                  Rose, Kenneth A.
;                  Stapleton, Genevieve
;       TITLE OF INVENTION: HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;                           SEQUENCES CODING THEREFOR AND USED THEREOF
;       NUMBER OF SEQUENCES: 45
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: NIXON & VANDERHYE P.C.
;            STREET: 1100 No. 6184350th Glebe Rd.  8th floor
;            CITY: Arlington
;            STATE: VA
;            COUNTRY: USA
;            ZIP: 22201-4741
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/270,751
;            FILING DATE: 17-Apr-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: PCT/GB95/02465
;            FILING DATE: 18-OCT-1995
;            APPLICATION NUMBER: GB 9421093.7
;            FILING DATE: 19-OCT-1994
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Wilson, Mary J.
;            REGISTRATION NUMBER: 32,955
;            REFERENCE/DOCKET NUMBER: 604-408
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 703-816-4000
;            TELEFAX: 703-816-4100
;   INFORMATION FOR SEQ ID NO: 24:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 14 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: <Unknown>
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-270-751-24

  Query Match             100.0%;  Score 33;  DB 2;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGWGVRQCLG 10


RESULT 36
US-09-270-751-25
; Sequence 25, Application US/09270751
; Patent No. 6184350
;  GENERAL INFORMATION:
;       APPLICANT: Lathe, Richard
;                  Rose, Kenneth A.
;                  Stapleton, Genevieve
;       TITLE OF INVENTION: HIPPOCAMPUS-ASSOCIATED PROTEINS; DNA
;                           SEQUENCES CODING THEREFOR AND USED THEREOF
;       NUMBER OF SEQUENCES: 45
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: NIXON & VANDERHYE P.C.
;            STREET: 1100 No. 6184350th Glebe Rd.  8th floor
;            CITY: Arlington
;            STATE: VA
;            COUNTRY: USA
;            ZIP: 22201-4741
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/270,751
;            FILING DATE: 17-Apr-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: PCT/GB95/02465
;            FILING DATE: 18-OCT-1995
;            APPLICATION NUMBER: GB 9421093.7
;            FILING DATE: 19-OCT-1994
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Wilson, Mary J.
;            REGISTRATION NUMBER: 32,955
;            REFERENCE/DOCKET NUMBER: 604-408
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 703-816-4000
;            TELEFAX: 703-816-4100
;   INFORMATION FOR SEQ ID NO: 25:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 14 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: <Unknown>
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-270-751-25

  Query Match             100.0%;  Score 33;  DB 2;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          1 FGYGVRACLG 10


RESULT 37
5204096-22
;Patent No. 5204096
;    APPLICANT: NEURATH, ALEXANDER R.;KENT, B.H.
;    TITLE OF INVENTION: PRE-S GENE CODED PEPTIDE HEPATITIS B
;IMMUNOGENS, VACCINES, DIAGNOSTICS, AND SYNTHETIC LIPID VESICLE
;CARRIERS
;    NUMBER OF SEQUENCES: 36
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/338,028
;      FILING DATE: 14-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 698,499
;      FILING DATE: 05-FEB-1985
;      APPLICATION NUMBER: 587,090
;      FILING DATE: 07-MAR-1984
;SEQ ID NO:22:
;      LENGTH: 15
5204096-22

  Query Match             100.0%;  Score 33;  DB 6;  Length 15;
  Best Local Similarity   40.0%;  Pred. No. 1.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          4 FLGGTTVCLG 13


RESULT 38
US-09-471-276-1257
; Sequence 1257, Application US/09471276
; Patent No. 6822072
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT: Duclert A.
```

```
;  APPLICANT: Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6822072
;  FILE REFERENCE: GENSET.025CP1
;  CURRENT APPLICATION NUMBER: US/09/471,276
;  CURRENT FILING DATE: 1999-12-21
;  EARLIER APPLICATION NUMBER: 09/057,719
;  EARLIER FILING DATE: 1998-04-09
;  EARLIER APPLICATION NUMBER: 09/069,047
;  EARLIER FILING DATE: 1998-04-28
;  EARLIER APPLICATION NUMBER: PCT/IB99/00712
;  EARLIER FILING DATE: 1999-04-09
;  NUMBER OF SEQ ID NOS: 1622
;  SOFTWARE: Patent.pm
; SEQ ID NO 1257
;   LENGTH: 16
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIGNAL
;   LOCATION: -14..-1
US-09-471-276-1257

  Query Match             100.0%;  Score 33;  DB 2;  Length 16;
  Best Local Similarity   40.0%;  Pred. No. 1.6e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          4 PFFGNSPCCG 13


RESULT 39
US-08-102-863-2
; Sequence 2, Application US/08102863
; Patent No. 5466590
;  GENERAL INFORMATION:
;    APPLICANT:  SARIASLANI, SIMA
;    TITLE OF INVENTION:  CONSTITUTIVE
;    TITLE OF INVENTION:  EXPRESSION OF P450SOY
;    TITLE OF INVENTION:  AND FERREDOXIN-SOY IN
;    TITLE OF INVENTION:  STREPTOMYCES
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E. I. DU PONT DE NEMOURS
;      ADDRESSEE:  AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  USA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0,
;      SOFTWARE:  Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/102,863
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/807,001
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GALLEGOS, R. THOMAS
;      REGISTRATION NUMBER:  32,692
;      REFERENCE/DOCKET NUMBER:  CR-9000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-7342
;      TELEFAX:  302-892-7949
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-102-863-2

  Query Match             100.0%;  Score 33;  DB 1;  Length 19;
  Best Local Similarity   40.0%;  Pred. No. 1.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          5 FGFGVHQCLG 14


RESULT 40
PCT-US92-10885-2
; Sequence 2, Application PC/TUS9210885
;  GENERAL INFORMATION:
;    APPLICANT:  SARIASLANI, SIMA
;    TITLE OF INVENTION:  CONSTITUTIVE
;    TITLE OF INVENTION:  EXPRESSION OF P450SOY
;    TITLE OF INVENTION:  AND FERREDOXIN-SOY IN
;    TITLE OF INVENTION:  STREPTOMYCES
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E. I. DU PONT DE NEMOURS
;      ADDRESSEE:  AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  USA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch,
;      MEDIUM TYPE:  1.0 MB
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  Macintosh System, 6.0
;      SOFTWARE:  Microsoft Word, 4.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/10885
;      FILING DATE:  19921216
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GALLEGOS, R. THOMAS
;      REGISTRATION NUMBER:  32,692
;      REFERENCE/DOCKET NUMBER:  CR-9000-A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-892-7342
;      TELEFAX:  302-892-7949
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 amino acids
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
PCT-US92-10885-2

  Query Match             100.0%;  Score 33;  DB 4;  Length 19;
  Best Local Similarity   40.0%;  Pred. No. 1.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          5 FGFGVHQCLG 14


RESULT 41
US-08-484-135-27
; Sequence 27, Application US/08484135
; Patent No. 5767078
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Johnson, Dana L
;    APPLICANT:  Zivin, Robert A
;    TITLE OF INVENTION:  AGONIST PEPTIDE DIMERS
;    NUMBER OF SEQUENCES:  93
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Frank S. DiGiglio
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A..
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,135
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9594
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-135-27

  Query Match             100.0%;  Score 33;  DB 1;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 1.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          8 FGPGTWDCTG 17


RESULT 42
US-08-484-635-208
; Sequence 208, Application US/08484635
; Patent No. 5773569
;  GENERAL INFORMATION:
;    APPLICANT:  Wrighton, Nicholas C.
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Chang, Ray S.
;    APPLICANT:  Kashyap, Arun K.
;    APPLICANT:  Jolliffe, Linda K.
;    APPLICANT:  Johnson, Dana
;    APPLICANT:  Mulcahy, Linda
;    TITLE OF INVENTION:  Compounds and Peptides That Bind to the
;    TITLE OF INVENTION:  Erythropoietin Receptor
;    NUMBER OF SEQUENCES:  259
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Street Tower
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,635
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/155,940
;      FILING DATE:  19-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  16528A-43-1-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  208:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-635-208

  Query Match             100.0%;  Score 33;  DB 1;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 1.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          8 FGPGTWDCTG 17


RESULT 43
US-08-484-631-208
; Sequence 208, Application US/08484631
; Patent No. 5830851
;  GENERAL INFORMATION:
;    APPLICANT:  Wrighton, Nicholas C.
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Chang, Ray S.
;    APPLICANT:  Kashyap, Arun K.
;    APPLICANT:  Jolliffe, Linda K.
;    APPLICANT:  Johnson, Dana
;    APPLICANT:  Mulcahy, Linda
;    TITLE OF INVENTION:  Compounds and Peptides That Bind to the
;    TITLE OF INVENTION:  Erythropoietin Receptor
;    NUMBER OF SEQUENCES:  259
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Street Tower
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,631
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/155,940
;      FILING DATE:  19-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  16528A-43-1-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
```

```
;  INFORMATION FOR SEQ ID NO:  208:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-631-208

  Query Match             100.0%;  Score 33;  DB 1;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 1.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          8 FGPGTWDCTG 17


RESULT 44
US-08-827-570-208
; Sequence 208, Application US/08827570
; Patent No. 5986047
;  GENERAL INFORMATION:
;    APPLICANT:  Wrighton, Nicholas C.
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Chang, Ray S.
;    APPLICANT:  Kashyap, Arun K.
;    APPLICANT:  Jolliffe, Linda K.
;    APPLICANT:  Johnson, Dana
;    APPLICANT:  Mulcahy, Linda
;    TITLE OF INVENTION:  Compounds and Peptides That Bind to the
;    TITLE OF INVENTION:  Erythropoietin Receptor
;    NUMBER OF SEQUENCES:  259
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Street Tower
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/827,570
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/484,635
;      FILING DATE:  07-JUN-1995
;      APPLICATION NUMBER:  US 08/155,940
;      FILING DATE:  19-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  16528A-43-1-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  208:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-827-570-208

  Query Match             100.0%;  Score 33;  DB 1;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 1.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          8 FGPGTWDCTG 17


RESULT 45
US-08-724-466B-6
; Sequence 6, Application US/08724466B
; Patent No. 6063606
;  GENERAL INFORMATION:
;    APPLICANT:  Petkovich, P. Martin, White, Jay A.,
;    APPLICANT:  Beckett, Barbara R., Jones, Glenville
;    TITLE OF INVENTION:  Retinoid Metabolizing Protein
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Blake, Cassels & Graydon
;      STREET:  Box 25, Commerce Court West
;      CITY:  Toronto
;      ZIP:  M5L 1A9
;      COUNTRY:  Canada
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3 1/2 inch, 1.4 Mb storage
;      COMPUTER:  COMPAQ, IBM PC compatible
;      OPERATING SYSTEM:  MS-DOS 5.1
;      SOFTWARE:  WORD PERFECT
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/724,466B
;      FILING DATE:  October 1, 1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/667,546
;      FILING DATE:  June 21, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hunt, John C.
;      REGISTRATION NUMBER:  36,424
;      REFERENCE/DOCKET NUMBER:  50767/00004
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 863-4344
;      TELEFAX:  (416) 863-2653
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-724-466B-6

  Query Match             100.0%;  Score 33;  DB 2;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 1.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGGGPRLCPG 11


RESULT 46
US-08-724-466B-7
; Sequence 7, Application US/08724466B
; Patent No. 6063606
;  GENERAL INFORMATION:
;    APPLICANT:  Petkovich, P. Martin, White, Jay A.,
;    APPLICANT:  Beckett, Barbara R., Jones, Glenville
;    TITLE OF INVENTION:  Retinoid Metabolizing Protein
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Blake, Cassels & Graydon
;      STREET:  Box 25, Commerce Court West
;      CITY:  Toronto
;      ZIP:  M5L 1A9
;      COUNTRY:  Canada
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3 1/2 inch, 1.4 Mb storage
```

```
;       COMPUTER:  COMPAQ, IBM PC compatible
;       OPERATING SYSTEM:  MS-DOS 5.1
;       SOFTWARE:  WORD PERFECT
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/724,466B
;       FILING DATE:  October 1, 1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/667,546
;       FILING DATE:  June 21, 1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Hunt, John C.
;       REGISTRATION NUMBER:  36,424
;       REFERENCE/DOCKET NUMBER:  50767/00004
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (416) 863-4344
;       TELEFAX:  (416) 863-2653
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  20 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-724-466B-7

  Query Match             100.0%;  Score 33;  DB 2;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 1.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FSGGARNCIG 11


RESULT 47
US-08-724-466B-8
; Sequence 8, Application US/08724466B
; Patent No. 6063606
;  GENERAL INFORMATION:
;    APPLICANT:  Petkovich, P. Martin, White, Jay A.,
;    APPLICANT:  Beckett, Barbara R., Jones, Glenville
;    TITLE OF INVENTION:  Retinoid Metabolizing Protein
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Blake, Cassels & Graydon
;      STREET:  Box 25, Commerce Court West
;      CITY:  Toronto
;      ZIP:  M5L 1A9
;      COUNTRY:  Canada
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3 1/2 inch, 1.4 Mb storage
;      COMPUTER:  COMPAQ, IBM PC compatible
;      OPERATING SYSTEM:  MS-DOS 5.1
;      SOFTWARE:  WORD PERFECT
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/724,466B
;      FILING DATE:  October 1, 1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/667,546
;      FILING DATE:  June 21, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hunt, John C.
;      REGISTRATION NUMBER:  36,424
;      REFERENCE/DOCKET NUMBER:  50767/00004
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 863-4344
;      TELEFAX:  (416) 863-2653
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-724-466B-8

  Query Match             100.0%;  Score 33;  DB 2;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 1.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FSGGARNCIG 11


RESULT 48
US-08-724-466B-9
; Sequence 9, Application US/08724466B
; Patent No. 6063606
;  GENERAL INFORMATION:
;    APPLICANT:  Petkovich, P. Martin, White, Jay A.,
;    APPLICANT:  Beckett, Barbara R., Jones, Glenville
;    TITLE OF INVENTION:  Retinoid Metabolizing Protein
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Blake, Cassels & Graydon
;      STREET:  Box 25, Commerce Court West
;      CITY:  Toronto
;      ZIP:  M5L 1A9
;      COUNTRY:  Canada
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3 1/2 inch, 1.4 Mb storage
;      COMPUTER:  COMPAQ, IBM PC compatible
;      OPERATING SYSTEM:  MS-DOS 5.1
;      SOFTWARE:  WORD PERFECT
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/724,466B
;      FILING DATE:  October 1, 1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/667,546
;      FILING DATE:  June 21, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hunt, John C.
;      REGISTRATION NUMBER:  36,424
;      REFERENCE/DOCKET NUMBER:  50767/00004
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 863-4344
;      TELEFAX:  (416) 863-2653
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-724-466B-9

  Query Match             100.0%;  Score 33;  DB 2;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 1.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGTGPRNCIG 11


RESULT 49
US-08-724-466B-10
; Sequence 10, Application US/08724466B
; Patent No. 6063606
;  GENERAL INFORMATION:
;    APPLICANT:  Petkovich, P. Martin, White, Jay A.,
;    APPLICANT:  Beckett, Barbara R., Jones, Glenville
;    TITLE OF INVENTION:  Retinoid Metabolizing Protein
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Blake, Cassels & Graydon
;      STREET:  Box 25, Commerce Court West
;      CITY:  Toronto
```

```
;       ZIP:  M5L 1A9
;       COUNTRY:  Canada
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette, 3 1/2 inch, 1.4 Mb storage
;       COMPUTER:  COMPAQ, IBM PC compatible
;       OPERATING SYSTEM:  MS-DOS 5.1
;       SOFTWARE:  WORD PERFECT
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/724,466B
;       FILING DATE:  October 1, 1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/667,546
;       FILING DATE:  June 21, 1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Hunt, John C.
;       REGISTRATION NUMBER:  36,424
;       REFERENCE/DOCKET NUMBER:  50767/00004
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (416) 863-4344
;       TELEFAX:  (416) 863-2653
;   INFORMATION FOR SEQ ID NO:  10:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  20 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-724-466B-10

  Query Match             100.0%;  Score 33;  DB 2;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 1.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FSGGSRNCIG 11


RESULT 50
US-08-882-164D-6
; Sequence 6, Application US/08882164D
; Patent No. 6306624
;  GENERAL INFORMATION:
;    APPLICANT:  Petkovich, P. Martin, White, Jay A.,
;    APPLICANT:  Beckett, Barbara R., Jones, Glenville
;    TITLE OF INVENTION:  Retinoid Metabolizing Protein
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Blake, Cassels & Graydon
;      STREET:  Box 25, Commerce Court West
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5L 1A9
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3 1/2 inch, 1.4 Mb storage
;      COMPUTER:  COMPAQ, IBM PC compatible
;      OPERATING SYSTEM:  MS-DOS 5.1
;      SOFTWARE:  WORD PERFECT
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/882,164D
;      FILING DATE:  June 25, 1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/667,546
;      FILING DATE:  June 21, 1996
;      APPLICATION NUMBER:  08/724,466
;      FILING DATE:  October 1, 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hunt, John C.
;      REGISTRATION NUMBER:  36,424
;      REFERENCE/DOCKET NUMBER:  50767/00010
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 863-4344
;      TELEFAX:  (416) 863-2653
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-882-164D-6

  Query Match             100.0%;  Score 33;  DB 2;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 1.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGGGPRLCPG 11


Search completed:  March  8, 2006, 11:25:27
Job time : 49 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2006, 11:20:19 ; Search time 38 Seconds
(without alignments)
25.320 Million cell updates/sec

Title: US-10-751-235-14
Perfect score: 33
Sequence: 1 FXXGXXXCXG 10

Scoring table: BLOSUM62DX:*
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	100.0	29	2	A56591	E75 steroid recept
2	33	100.0	35	2	A82520	hypothetical prote
3	33	100.0	47	2	I48943	cellular disintegr
4	33	100.0	50	2	D72804	gp38 protein - Myc
5	33	100.0	63	2	S09643	probable cytochrom
6	33	100.0	66	2	A83575	hypothetical prote
7	33	100.0	70	2	T06525	cytochrome P450 ho
8	33	100.0	71	2	D81003	hypothetical prote
9	33	100.0	71	2	D82025	probable integral
10	33	100.0	74	2	T17791	hypothetical prote
11	33	100.0	78	2	T01646	apoptotic cell dea
12	33	100.0	80	2	S38926	hypothetical prote
13	33	100.0	87	2	T17723	hypothetical prote
14	33	100.0	89	2	T06524	probable cytochrom
15	33	100.0	96	2	AE2692	hypothetical prote
16	33	100.0	103	2	A75635	cytochrome P450-re
17	33	100.0	112	2	B82914	hypothetical prote
18	33	100.0	113	2	A54437	apoptotic cell dea
19	33	100.0	113	2	C54437	apoptotic cell dea
20	33	100.0	113	2	I49285	defender against a
21	33	100.0	114	2	T03016	probable apoptotic
22	33	100.0	115	2	S71269	apoptotic cell dea
23	33	100.0	115	2	F86446	probable defender
24	33	100.0	119	2	T17016	defender against c
25	33	100.0	119	2	T39694	hypothetical prote
26	33	100.0	119	2	H87152	hypothetical prote
27	33	100.0	122	2	T39039	probable oligosacc
28	33	100.0	122	2	T38989	hypothetical prote
29	33	100.0	124	2	T05467	hypothetical prote
30	33	100.0	126	2	I46935	aldehyde dehydroge
31	33	100.0	132	2	T28875	hypothetical prote
32	33	100.0	133	2	S61662	dolichyl-diphospho
33	33	100.0	136	2	F69870	general stress pro
34	33	100.0	137	2	S22388	phospholipase A2 (
35	33	100.0	137	2	A95297	hypothetical prote
36	33	100.0	139	2	G85060	hypothetical prote
37	33	100.0	139	2	E85060	hypothetical prote
38	33	100.0	141	2	T05465	hypothetical prote
39	33	100.0	141	2	H85060	hypothetical prote
40	33	100.0	141	2	T26567	hypothetical prote
41	33	100.0	142	2	T36312	hypothetical prote
42	33	100.0	143	2	T05466	hypothetical prote
43	33	100.0	144	2	PC4428	cytochrome P450 4C
44	33	100.0	145	2	S20227	cytochrome P450 2C
45	33	100.0	147	2	S36442	reverse transcript
46	33	100.0	148	2	G82223	PilB-related prote
47	33	100.0	157	2	S30389	hypothetical prote
48	33	100.0	160	2	F84769	defender against c
49	33	100.0	162	2	PC4262	cytochrome P450 mo
50	33	100.0	169	2	I55299	cytochrome P450 PB
51	33	100.0	173	2	S20689	hypothetical 19.4K
52	33	100.0	175	2	C81349	thioredoxin peroxi
53	33	100.0	176	2	F96824	hypothetical prote
54	33	100.0	178	2	AF0544	probable fimbrial
55	33	100.0	181	2	D69084	TATA-binding trans
56	33	100.0	183	2	C64363	TATA-binding trans
57	33	100.0	183	2	E69296	transcription init
58	33	100.0	186	2	T44922	TATA box binding p
59	33	100.0	186	2	T08255	TATA-binding trans
60	33	100.0	186	2	A84375	transcription init
61	33	100.0	186	2	T08331	TATA-binding trans
62	33	100.0	189	2	S30172	mercury-binding pr
63	33	100.0	190	2	JC4514	TATA-binding prote
64	33	100.0	191	2	T47230	transcription fact
65	33	100.0	191	2	A54275	TATA-binding prote
66	33	100.0	191	2	E75072	transcription init
67	33	100.0	191	2	D71093	probable TATA-bind
68	33	100.0	200	1	TWPO2D	transcription init
69	33	100.0	200	2	S61088	transcription init
70	33	100.0	200	2	S21140	transcription init
71	33	100.0	200	2	S10946	transcription init
72	33	100.0	200	2	S10945	transcription init
73	33	100.0	201	2	S30216	transcription init
74	33	100.0	205	2	T27278	hypothetical prote
75	33	100.0	209	2	B95083	thiamin-phosphate
76	33	100.0	209	2	F97950	thiamine-phosphate
77	33	100.0	213	2	E64030	probable holocytoc
78	33	100.0	215	2	T03386	transcription init
79	33	100.0	219	2	T08330	hypothetical prote
80	33	100.0	220	2	S48781	L1 protein - human
81	33	100.0	221	2	S48782	L1 protein - human
82	33	100.0	221	2	G83604	conserved hypothet
83	33	100.0	222	2	S48786	L1 protein - human
84	33	100.0	222	2	H64495	cobalamin biosynth
85	33	100.0	224	2	S37740	transcription fact
86	33	100.0	224	2	S48779	L1 protein - human
87	33	100.0	224	2	S48784	L1 protein - human
88	33	100.0	225	2	S48780	L1 protein - human
89	33	100.0	226	1	JQ1570	major surface anti
90	33	100.0	226	1	JQ1571	major surface anti
91	33	100.0	226	1	JQ1573	major surface anti
92	33	100.0	226	1	JQ1574	major surface anti
93	33	100.0	226	1	JQ1576	major surface anti
94	33	100.0	226	1	JQ1577	major surface anti
95	33	100.0	226	1	JQ1578	major surface anti
96	33	100.0	226	1	JQ1579	major surface anti
97	33	100.0	226	1	JQ1580	major surface anti
98	33	100.0	226	1	JQ1581	major surface anti
99	33	100.0	226	1	SAVLAD	major surface anti
100	33	100.0	226	1	SAVLAR	major surface anti

ALIGNMENTS

RESULT 1
A56591
E75 steroid receptor homolog Ms1 - tobacco hornworm (fragment)
C;Species: Manduca sexta (tobacco hornworm)
C;Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 09-Jul-2004
C;Accession: A56591
R;Segraves, W.A.; Woldin, C.
Insect Biochem. Mol. Biol. 23, 91-97, 1993
A;Title: The E75 gene of Manduca sexta and comparison with its Drosophila homolog.
A;Reference number: A56591; MUID:93250864; PMID:8485520
A;Accession: A56591
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-29 <SEG>
A;Cross-references: UNIPROT:Q08893; UNIPARC:UPI000016BFF9; GB:S60782; NID:g385887; PIDN:
A;Note: sequence extracted from NCBI backbone (NCBIN:132107, NCBIP:132113)
C;Keywords: steroid hormone receptor

Query Match 100.0%; Score 33; DB 2; Length 29;
Best Local Similarity 40.0%; Pred. No. 89;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|::|:::|:|
Db 18 FHYGVHSCEG 27

RESULT 2
A82520
hypothetical protein XF2749 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: A82520
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: A82520
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-35 <SIM>
A;Cross-references: UNIPROT:Q9P9X1; UNIPARC:UPI00000C2B3B; GB:AE004081; GB:AE003849; NID
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF2749

Query Match 100.0%; Score 33; DB 2; Length 35;
Best Local Similarity 40.0%; Pred. No. 1e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|::|:::|:|
Db 5 FDAGAAICVG 14

RESULT 3
I48943
cellular disintegrin-related protein 16-2 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 15-Mar-1996 #sequence_revision 15-Mar-1996 #text_change 05-Oct-2004
C;Accession: I48943
R;Weskamp, G.; Blobel, C.P.
Proc. Natl. Acad. Sci. U.S.A. 91, 2748-2751, 1994
A;Title: A new family of cellular proteins related to snake venom disintegrins.
A;Reference number: A53476; MUID:94195820; PMID:8146185
A;Accession: I48943
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-47 <RES>
A;Cross-references: UNIPROT:Q61072; UNIPARC:UPI000016CAB9; EMBL:U06145; NID:g487138; PIDN

Query Match 100.0%; Score 33; DB 2; Length 47;
Best Local Similarity 40.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|::|:::|:|
Db 35 FLPGGSMCRG 44

RESULT 4
D72804
gp38 protein - Mycobacterium phage D29
C;Species: Mycobacterium phage D29
C;Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 09-Jul-2004
C;Accession: D72804
R;Ford, M.E.; Sarkis, G.J.; Belanger, A.E.; Hendrix, R.W.; Hatfull, G.F.
J. Mol. Biol. 279, 143-164, 1998
A;Title: Genome structure of mycobacteriophage D29: Implications for phage evolution.
A;Reference number: A72800; MUID:98300335; PMID:9636706
A;Accession: D72804
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-50 <FOR>
A;Cross-references: UNIPROT:O64229; UNIPARC:UPI000013856F; GB:AF022214; NID:g3172250; PI
C;Genetics:
A;Gene: 38

Query Match 100.0%; Score 33; DB 2; Length 50;
Best Local Similarity 40.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|::|:::|:|
Db 13 FALGLTACDG 22

RESULT 5
S09643
probable cytochrome P450 - Neurospora crassa (fragment)
C;Species: Neurospora crassa
C;Date: 21-Nov-1993 #sequence_revision 17-Apr-1998 #text_change 05-Oct-2004
C;Accession: S09643
R;Attar, R.M.; Grotewold, E.; Taccioli, G.E.; Aisemberg, G.O.; Torres, H.N.; Judewicz, N
Nucleic Acids Res. 17, 7535-7536, 1989
A;Title: A cycloheximide-inducible gene of Neurospora crassa belongs to the cytochrome P
A;Reference number: S09643; MUID:90016828; PMID:2529480
A;Accession: S09643
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-63 <ATT>
A;Cross-references: UNIPARC:UPI000017A254; EMBL:X15033
C;Genetics:
A;Gene: CYP54
A;Introns: 11/2
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein
F;8/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 100.0%; Score 33; DB 2; Length 63;
Best Local Similarity 40.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|::|:::|:|
Db 1 FGYGRHACPG 10

RESULT 6
A83575
hypothetical protein PA0553 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: A83575
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: A83575
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-66 <STO>
A;Cross-references: UNIPROT:Q9I5Y3; UNIPARC:UPI00000C50AF; GB:AE004492; GB:AE004091; NID
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA0553

Query Match 100.0%; Score 33; DB 2; Length 66;
Best Local Similarity 40.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|::|:::|:|
Db 11 FCLGLAACSG 20

RESULT 7
T06525
cytochrome P450 homolog - garden pea (fragment)
C;Species: Pisum sativum (garden pea)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T06525
R;Frank, M.R.; Deyneka, J.M.; Schuler, M.A.
Plant Physiol. 110, 1035-1046, 1996
A;Title: Cloning of wound-induced cytochrome P450 mono-oxygenases expressed in pea.
A;Reference number: Z15734; MUID:96417083; PMID:8819874
A;Accession: T06525
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-70 <FRA>
A;Cross-references: UNIPROT:Q40986; UNIPARC:UPI000009CF11; EMBL:U29335; NID:g894156; PID
C;Superfamily: Candida cytochrome P450 52A3; cytochrome P450 homology

Query Match 100.0%; Score 33; DB 2; Length 70;
Best Local Similarity 40.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|::|:::|:|
Db 6 FQAGPRVCLG 15

RESULT 8
D81003
hypothetical protein NMB2131 [imported] - Neisseria meningitidis (strain MC58 serogroup
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: D81003
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ver
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: D81003
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-71 <TET>
A;Cross-references: UNIPROT:Q9JXB9; UNIPARC:UPI00000C487B; GB:AE002561; GB:AE002098; NID
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB2131

Query Match 100.0%; Score 33; DB 2; Length 71;
Best Local Similarity 40.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|::|:::|:|
Db 2 FKLGVYTCLG 11

RESULT 9
D82025
probable integral membrane protein NMA0300 [imported] - Neisseria meningitidis (strain Z2
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: D82025
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morell
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria menigitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: D82025
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-71 <PAR>
A;Cross-references: UNIPROT:Q9JWM3; UNIPARC:UPI00000C4944; GB:AL162752; GB:AL157959; NID
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA0300

Query Match 100.0%; Score 33; DB 2; Length 71;
Best Local Similarity 40.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FXXGXXXCXG 10
|::|:::|:|
Db 2 FKLGVYACLG 11

RESULT 10
T17791
hypothetical protein a294R - Chlorella virus PBCV-1
C;Species: Chlorella virus PBCV-1
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17791
R;Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z18806
A;Accession: T17791
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-74 <GRA>
A;Cross-references: UNIPROT:Q84610; UNIPARC:UPI00000F54FF; EMBL:U42580; NID:g4028896; PII
A;Experimental source: specific host Chlorella strain NC64A
C;Genetics:
A;Note: a294R

Query Match 100.0%; Score 33; DB 2; Length 74;
Best Local Similarity 40.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         16 FFKGASTCEG 25


RESULT 11
T01646
apoptotic cell death regulator DAD1 homolog - maize (fragment)
N;Alternate names: apoptosis suppressor; defender against cell death 1 protein
C;Species: Zea mays (maize)
C;Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 31-Dec-2004
C;Accession: T01646
R;Finkelstein, D.B.; Drew, M.C.; Jordan, W.R.; Wing, R.A.; Mullet, J.E.; Morgan, P.W.
submitted to the EMBL Data Library, March 1998
A;Description: Apoptotic gene discovery in maize.
A;Reference number: Z14356
A;Accession: T01646
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-78 <FIN>
A;Cross-references: UNIPARC:UPI0000178A85; EMBL:AF055909; NID:g3264593; PIDN:AAC24568.1;
A;Experimental source: strain TX5855
C;Genetics:
A;Gene: dad1
C;Superfamily: Defender against cell death DAD
C;Keywords: apoptosis

  Query Match             100.0%;  Score 33;  DB 2;  Length 78;
  Best Local Similarity   40.0%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         23 FLSGVLSCIG 32


RESULT 12
S38926
hypothetical protein 14 - phage phi-C31
C;Species: phage phi-C31
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: S38926
R;Hartley, N.M.; Murphy, G.O.; Bruton, C.J.; Chater, K.F.
submitted to the EMBL Data Library, November 1993
A;Reference number: S38912
A;Accession: S38926
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-80 <HAR>
A;Cross-references: UNIPROT:Q38035; UNIPARC:UPI000009AEB0; EMBL:X76288; NID:g432610; PID
C;Genetics:
A;Start codon: GTG

  Query Match             100.0%;  Score 33;  DB 2;  Length 80;
  Best Local Similarity   40.0%;  Pred. No. 2.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         16 FPSGPYTCEG 25


RESULT 13
T17723
hypothetical protein a232R - Chlorella virus PBCV-1
C;Species: Chlorella virus PBCV-1
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17723
R;Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z18806
A;Accession: T17723
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-87 <GRA>
A;Cross-references: UNIPROT:Q84552; UNIPARC:UPI00000F617B; EMBL:U42580; NID:g4028896; PII
A;Experimental source: specific host Chlorella strain NC64A
C;Genetics:
A;Note: a232R

  Query Match             100.0%;  Score 33;  DB 2;  Length 87;
  Best Local Similarity   40.0%;  Pred. No. 2.2e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         26 FTEGSEKCSG 35


RESULT 14
T06524
probable cytochrome P450 - garden pea (fragment)
N;Contains: oxidoreductase (EC 1.-.-.-)
C;Species: Pisum sativum (garden pea)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T06524
R;Frank, M.R.; Deyneka, J.M.; Schuler, M.A.
Plant Physiol. 110, 1035-1046, 1996
A;Title: Cloning of wound-induced cytochrome P450 mono-oxygenases expressed in pea.
A;Reference number: Z15734; MUID:96417083; PMID:8819874
A;Accession: T06524
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-89 <FRA>
A;Cross-references: UNIPROT:Q40985; UNIPARC:UPI000009FA6B; EMBL:U29334; NID:g894154; PIDN
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C;Keywords: chromoprotein; electron transfer; heme; iron; monooxygenase; oxidoreductase

  Query Match             100.0%;  Score 33;  DB 2;  Length 89;
  Best Local Similarity   40.0%;  Pred. No. 2.3e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         25 FRSGRRGCPG 34


RESULT 15
AE2692
hypothetical protein Atu0941 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AE2692
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AE2692
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-96 <KUR>
A;Cross-references: UNIPROT:Q8UGU4; UNIPARC:UPI00000D19AD; GB:AE008688; PIDN:AAL41955.1;
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu0941
A;Map position: circular chromosome

  Query Match             100.0%;  Score 33;  DB 2;  Length 96;
  Best Local Similarity   40.0%;  Pred. No. 2.4e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         18 FLIGVVLCFG 27


RESULT 16
A75635
cytochrome P450-related protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: A75635
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: A75635
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-103 <WHI>
A;Cross-references: UNIPROT:Q9RZI2; UNIPARC:UPI00000D3B6F; GB:AE001827; NID:g6460959; PI
A;Experimental source: strain R1
C;Genetics:
A;Gene: DRC0001
A;Map position: plasmid
A;Genome: plasmid
A;Note: plasmid CP1
C;Superfamily: Deinococcus radiodurans cytochrome P450-related protein DRC0001

  Query Match             100.0%;  Score 33;  DB 2;  Length 103;
  Best Local Similarity   40.0%;  Pred. No. 2.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         63 FGKGEHACLG 72


RESULT 17
B82914
hypothetical protein UU225.1 [imported] - Ureaplasma urealyticum
C;Species: Ureaplasma urealyticum
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: B82914
R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
A;Reference number: A82870
A;Accession: B82914
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-112 <GLA>
A;Cross-references: UNIPARC:UPI00000C1BDB; GB:AE002123; GB:AF222894; NID:g6899229; PIDN:
A;Experimental source: serovar 3; biovar 1
C;Genetics:
A;Gene: UU225.1
A;Genetic code: SGC3

  Query Match             100.0%;  Score 33;  DB 2;  Length 112;
  Best Local Similarity   40.0%;  Pred. No. 2.7e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         76 FVFGILGCLG 85


RESULT 18
A54437
apoptotic cell death regulator DAD1 - human
C;Species: Homo sapiens (man)
C;Date: 13-Sep-1994 #sequence_revision 18-Nov-1994 #text_change 31-Dec-2004
C;Accession: A54437
R;Nakashima, T.; Sekiguchi, T.; Kuraoka, A.; Fukushima, K.; Shibata, Y.; Komiyama, S.; Ni
Mol. Cell. Biol. 13, 6367-6374, 1993
A;Title: Molecular cloning of a human cDNA encoding a novel protein, DAD1, whose defect c
A;Reference number: A54437; MUID:94019310; PMID:8413235
A;Accession: A54437
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-113 <NAK>
A;Cross-references: UNIPROT:P61803; UNIPARC:UPI00000040DF; GB:D15057; NID:g493244; PIDN:E
A;Experimental source: transformed Raji cells
A;Note: sequence extracted from NCBI backbone (NCBIN:137990, NCBIP:137993)
C;Superfamily: Defender against cell death DAD

  Query Match             100.0%;  Score 33;  DB 2;  Length 113;
  Best Local Similarity   40.0%;  Pred. No. 2.7e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         56 FLSGFISCVG 65


RESULT 19
C54437
apoptotic cell death regulator DAD1 - golden hamster
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 31-Dec-2004
C;Accession: C54437
R;Nakashima, T.; Sekiguchi, T.; Kuraoka, A.; Fukushima, K.; Shibata, Y.; Komiyama, S.; Ni
Mol. Cell. Biol. 13, 6367-6374, 1993
A;Title: Molecular cloning of a human cDNA encoding a novel protein, DAD1, whose defect c
A;Reference number: A54437; MUID:94019310; PMID:8413235
A;Accession: C54437
A;Molecule type: mRNA
A;Residues: 1-113 <NAK>
A;Cross-references: UNIPROT:P61806; UNIPARC:UPI00000040DF; GB:D15058; NID:g493235; PIDN:E
A;Experimental source: BHK21 cell line
C;Superfamily: Defender against cell death DAD

  Query Match             100.0%;  Score 33;  DB 2;  Length 113;
  Best Local Similarity   40.0%;  Pred. No. 2.7e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         56 FLSGFISCVG 65


RESULT 20
I49285
defender against apoptotic cell death protein 1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 31-Dec-2004
C;Accession: I49285; S65658
R;Apte, S.S.; Mattei, M.G.; Seldin, M.F.; Olsen, B.R.
FEBS Lett. 363, 304-306, 1995
A;Title: The highly conserved defender against the death 1 (DAD1) gene maps to human chr
A;Reference number: I49285; MUID:95255568; PMID:7737422
A;Accession: I49285
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-113 <RES>
A;Cross-references: UNIPROT:P61804; UNIPARC:UPI0000047517; EMBL:U22107; NID:g849126; PIDN
A;Accession: S65658
A;Molecule type: mRNA
A;Residues: 1-113 <APT>
A;Cross-references: UNIPARC:UPI0000047517; GB:U22107; NID:g849126; PIDN:AAA85855.1; PID:g
C;Genetics:
A;Gene: DAD1
A;Map position: 14
C;Superfamily: Defender against cell death DAD
```

```
  Query Match             100.0%;  Score 33;  DB 2;  Length 113;
  Best Local Similarity   40.0%;  Pred. No. 2.7e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         56 FLSGFISCVG 65


RESULT 21
T03016
probable apoptotic cell death regulator DAD1 - rice
C;Species: Oryza sativa (rice)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 31-Dec-2004
C;Accession: T03016
R;Tanaka, Y.; Makishima, T.; Sasabe, M.; Ichinose, Y.; Shiraishi, T.; Nishimoto, T.; Yam
Plant Cell Physiol. 38, 383-397, 1997
A;Title: Dad-1, a putative programmed cell death gene in rice.
A;Reference number: Z14819
A;Accession: T03016
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-114 <TAN>
A;Cross-references: UNIPROT:O50070; UNIPARC:UPI0000128DE9; EMBL:D89726; PIDN:BAA24072.1
A;Experimental source: cultivar Nipponbare
C;Genetics:
A;Gene: DAD-1
C;Superfamily: Defender against cell death DAD

  Query Match             100.0%;  Score 33;  DB 2;  Length 114;
  Best Local Similarity   40.0%;  Pred. No. 2.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         58 FLSGVLSCIG 67


RESULT 22
S71269
apoptotic cell death regulator DAD1 homolog - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change 31-Dec-2004
C;Accession: S71269
R;Cooke, M.R.
submitted to the EMBL Data Library, February 1996
A;Reference number: S71269
A;Accession: S71269
A;Molecule type: mRNA
A;Residues: 1-115 <COO>
A;Cross-references: UNIPROT:Q39080; UNIPARC:UPI0000128DE1; EMBL:X95585; NID:g1184192; PI
C;Superfamily: Defender against cell death DAD
C;Keywords: apoptosis

  Query Match             100.0%;  Score 33;  DB 2;  Length 115;
  Best Local Similarity   40.0%;  Pred. No. 2.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         59 FLSGVLSCIG 68


RESULT 23
F86446
probable defender against cell death protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2004
C;Accession: F86446
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: F86446
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-115 <STO>
A;Cross-references: UNIPROT:Q39080; UNIPARC:UPI0000162F4A; GB:AE005172; NID:g10801366; PI
C;Genetics:
A;Map position: 1
C;Superfamily: Defender against cell death DAD

  Query Match             100.0%;  Score 33;  DB 2;  Length 115;
  Best Local Similarity   40.0%;  Pred. No. 2.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         59 FLSGVLSCIG 68


RESULT 24
T17016
defender against cell death 1 homolog - apple tree
C;Species: Malus domestica (apple tree)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Dec-2004
C;Accession: T17016
R;Dong, Y.H.; Atkinson, R.G.; Morris, B.A.; Gardner, R.C.
submitted to the EMBL Data Library, August 1996
A;Description: An apple mRNA with high homolgy to the human DAD-1 gene.
A;Reference number: Z18653
A;Accession: T17016
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-119 <DON>
A;Cross-references: UNIPROT:O24060; UNIPARC:UPI0000128DE8; EMBL:U68560; NID:g1546826; PII
A;Experimental source: strain Granny Smith
C;Genetics:
A;Gene: DAD-1
C;Superfamily: Defender against cell death DAD
C;Keywords: apoptosis

  Query Match             100.0%;  Score 33;  DB 2;  Length 119;
  Best Local Similarity   40.0%;  Pred. No. 2.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         63 FLSGVLSCIG 72


RESULT 25
T39694
hypothetical protein SPAC17A3.02 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T39694
R;Wood, V.; Skelton, J.; Churcher, C.M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, July 1999
A;Reference number: Z21870
A;Accession: T39694
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-119 <WOO>
A;Cross-references: UNIPROT:Q9UUF4; UNIPARC:UPI000006A199; EMBL:AL109652; PIDN:CAB51761.
A;Experimental source: strain 972h-; cosmid c17A3
C;Genetics:
A;Gene: SPDB:SPAC17A3.02
A;Map position: 2
```

```
A;Introns: 23/1; 28/1; 59/2; 66/1

  Query Match            100.0%;  Score 33;  DB 2;  Length 119;
  Best Local Similarity   40.0%;  Pred. No. 2.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         44 FVPGCLLCGG 53


RESULT 26
H87152
hypothetical protein [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: H87152
R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A;Title: Massive gene decay in the leprosy bacillus.
A;Reference number: A86909; MUID:21128732; PMID:11234002
A;Accession: H87152
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-119 <STO>
A;Cross-references: UNIPROT:Q9CBH8; UNIPARC:UPI00000C6E17; GB:AL450380; NID:g13093604; P
C;Genetics:
A;Gene: ML1949

  Query Match            100.0%;  Score 33;  DB 2;  Length 119;
  Best Local Similarity   40.0%;  Pred. No. 2.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FWVGVRGCGG 11


RESULT 27
T39039
probable oligosaccharyl transferase epsilon subunitdefender against cell death homolog -
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Dec-2004
C;Accession: T39039
R;Gentles, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A;Reference number: Z21797
A;Accession: T39039
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-122 <GEN>
A;Cross-references: UNIPROT:O14238; UNIPARC:UPI0000130EE7; EMBL:Z98981; PIDN:CAB11729.1;
A;Experimental source: strain 972h-; cosmid c6F6
C;Genetics:
A;Gene: SPDB:SPAC6F6.05
A;Map position: 1
A;Introns: 101/2
C;Superfamily: Defender against cell death DAD

  Query Match            100.0%;  Score 33;  DB 2;  Length 122;
  Best Local Similarity   40.0%;  Pred. No. 2.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         58 FLSGFISCVG 67


RESULT 28
T38989
hypothetical protein SPAC630.11 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T38989
R;McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Zimmermann, W.; Wambutt, R.
submitted to the EMBL Data Library, August 1999
A;Reference number: Z21822
A;Accession: T38989
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-122 <MCD>
A;Cross-references: UNIPROT:Q9UUH1; UNIPARC:UPI0000069E76; EMBL:AL109832; PIDN:CAB52733.
A;Experimental source: strain 972h-; cosmid c630
C;Genetics:
A;Gene: SPDB:SPAC630.11
A;Map position: 1
A;Introns: 6/2; 12/3; 37/1; 59/2; 78/1
C;Superfamily: Saccharomyces probable membrane protein YJR044c

  Query Match            100.0%;  Score 33;  DB 2;  Length 122;
  Best Local Similarity   40.0%;  Pred. No. 2.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         67 FTFGATICTG 76


RESULT 29
T05467
hypothetical protein T8O5.60 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05467
R;Bevan, M.; Wedler, H.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schuelle
submitted to the Protein Sequence Database, February 1998
A;Reference number: Z15417
A;Accession: T05467
A;Molecule type: DNA
A;Residues: 1-124 <BEV>
A;Cross-references: UNIPROT:O49708; UNIPARC:UPI00000A83C8; EMBL:AL021890
A;Experimental source: cultivar Columbia; BAC clone T8O5
C;Genetics:
A;Map position: 4
A;Introns: 40/2; 92/3
A;Note: T8O5.60
C;Superfamily: hypothetical protein YCL033c

  Query Match            100.0%;  Score 33;  DB 2;  Length 124;
  Best Local Similarity   40.0%;  Pred. No. 3e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         51 FDDGIYSCAG 60


RESULT 30
I46935
aldehyde dehydrogenase class x - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 09-Jul-2004
C;Accession: I46935
R;Algar, E.M.; Cheung, B.; Hayes, J.; Holmes, R.S.; Beacham, I.R.
Adv. Exp. Med. Biol. 328, 153-157, 1993
A;Title: Bovine corneal aldehyde dehydrogenases: evidence for multiple gene products (ALI
A;Reference number: I46935; MUID:93263009; PMID:8493893
A;Accession: I46935
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-126 <ALG>
A;Cross-references: UNIPROT:P52476; UNIPARC:UPI0000129261; GB:S61045; NID:g300403; PIDN:
C;Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
```

```
  Query Match             100.0%;  Score 33;  DB 2;  Length 126;
  Best Local Similarity   40.0%;  Pred. No. 3e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         24 FSMGQCCCPG 33


RESULT 31
T28875
hypothetical protein R04E5.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
C;Accession: T28875
R;Miller, N.
submitted to the EMBL Data Library, December 1995
A;Description: The sequence of C. elegans cosmid R04E5.
A;Reference number: Z20535
A;Accession: T28875
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-132 <MIL>
A;Cross-references: UNIPARC:UPI0000179CD0; EMBL:U41538; PIDN:AAC48178.1; GSPDB:GN00028;
A;Experimental source: strain Bristol N2; clone R04E5
C;Genetics:
A;Gene: CESP:R04E5.1
A;Map position: X
A;Introns: 31/3; 72/3; 102/1; 125/3

  Query Match             100.0%;  Score 33;  DB 2;  Length 132;
  Best Local Similarity   40.0%;  Pred. No. 3.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         77 FECGPWTCFG 86


RESULT 32
S61662
dolichyl-diphosphooligosaccharide-protein glycotransferase (EC 2.4.1.119) epsilon chain
N;Alternate names: protein O3211; protein YOR103c; protein YOR3211c
C;Species: Saccharomyces cerevisiae
C;Date: 09-Mar-1996 #sequence_revision 12-Apr-1996 #text_change 31-Dec-2004
C;Accession: S61662; S62142; S62149; S66988
R;Benes, V.; Andrade, M.A.; Rechmann, S.; Teodoru, C.; Banrevi, A.; Sander, C.; Valencia
submitted to the EMBL Data Library, December 1995
A;Description: Nucleotide sequence and analysis of a 130 kb fragment of yeast chromosome
A;Reference number: S61643
A;Accession: S61662
A;Molecule type: DNA
A;Residues: 1-133 <BEN>
A;Cross-references: UNIPROT:P46964; UNIPARC:UPI0000168417; EMBL:X94335; NID:g1262139; PI
R;Silberstein, S.; Collins, P.G.; Kelleher, D.J.; Gilmore, R.
J. Cell Biol. 131, 371-383, 1995
A;Title: The essential OST2 gene encodes the 16-kD subunit of the yeast oligosaccharyltr
A;Reference number: S62142; MUID:96017708; PMID:7593165
A;Accession: S62142
A;Molecule type: DNA
A;Residues: 4-133 <SIL>
A;Cross-references: UNIPARC:UPI0000168A33; EMBL:U32307; NID:g1041719; PIDN:AAC49086.1; P
A;Accession: S62149
A;Molecule type: protein
A;Residues: 5-31,36-43 <SIW>
A;Cross-references: UNIPARC:UPI0000178A86
R;Voss, H.; Benes, V.; Rechmann, S.; Teodoru, C.; Schwager, C.; Paces, V.; Ansorge, W.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S66965
A;Accession: S66988
A;Molecule type: DNA
A;Residues: 1-133 <VOS>
A;Cross-references: UNIPARC:UPI0000168417; EMBL:Z75010; NID:g1420282; PIDN:CAA99300.1; PI
A;Experimental source: strain S288C.
C;Genetics:
A;Gene: SGD:OST2; OST2
A;Cross-references: SGD:S0005629
A;Map position: 15R
C;Superfamily: Defender against cell death DAD
C;Keywords: glycosyltransferase; hexosyltransferase; transmembrane protein
F;55-71/Domain: transmembrane #status predicted <TM1>
F;79-95/Domain: transmembrane #status predicted <TM2>
F;113-129/Domain: transmembrane #status predicted <TM3>

  Query Match             100.0%;  Score 33;  DB 2;  Length 133;
  Best Local Similarity   40.0%;  Pred. No. 3.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         80 FLAGFIICVG 89


RESULT 33
F69870
general stress protein homolog ykzA - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: F69870
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berterc
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallerc
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: F69870
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-136 <KUN>
A;Cross-references: UNIPROT:P80242; UNIPARC:UPI00000602B3; GB:Z99110; GB:AL009126; NID:g
A;Experimental source: strain 168
C;Genetics:
A;Gene: ykzA
C;Superfamily: hypothetical protein yklA

  Query Match             100.0%;  Score 33;  DB 2;  Length 136;
  Best Local Similarity   40.0%;  Pred. No. 3.2e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         48 FAAGYAACFG 57


RESULT 34
S22388
phospholipase A2 (EC 3.1.1.4) ammodytin I2 precursor - western sand viper
C;Species: Vipera ammodytes ammodytes (western sand viper)
C;Date: 07-Apr-1994 #sequence_revision 19-May-1994 #text_change 05-Oct-2004
C;Accession: S22388; S36686; S36685
R;Krizaj, I.; Liang, N.S.; Pungercar, J.; Strukelj, B.; Ritonja, A.; Gubensek, F.
Eur. J. Biochem. 204, 1057-1062, 1992
A;Title: Amino acid and cDNA sequences of a neutral phospholipase A(2) from the long-nos
A;Reference number: S22388; MUID:92201190; PMID:1551386
A;Accession: S22388
```

```
A;Molecule type: mRNA
A;Residues: 1-137 <KRI>
A;Cross-references: UNIPROT:P34180; UNIPARC:UPI0000131159; EMBL:X56878; NID:g5702035; PI
A;Accession: S36686
A;Status: preliminary
A;Molecule type: protein
A;Residues: 17-137 <KR2>
A;Cross-references: UNIPARC:UPI0000175885
R;Gubensek, F.
submitted to the EMBL Data Library, January 1991
A;Reference number: S36685
A;Accession: S36685
A;Molecule type: mRNA
A;Residues: 1-131,'S',133-137 <GUB>
A;Cross-references: UNIPARC:UPI0000175886; EMBL:X56878; NID:g64450; PIDN:CAA40200.1; PID
C;Superfamily: Phospholipase A2
C;Keywords: calcium; calcium binding; carboxylic ester hydrolase; venom
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-137/Product: ammodytin I2 #status predicted <MAT>
F;43,45,47,64/Binding site: calcium (Tyr, Gly, Gly, Asp) #status predicted
F;63,105/Active site: His, Asp #status predicted

  Query Match             100.0%;  Score 33;  DB 2;  Length 137;
  Best Local Similarity   40.0%;  Pred. No. 3.2e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         84 FENGDIVCGG 93


RESULT 35
A95297
hypothetical protein SMa0537 [imported] - Sinorhizobium meliloti (strain 1021) magaplasm
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: A95297
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: A95297
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-137 <KUR>
A;Cross-references: UNIPROT:Q930B7; UNIPARC:UPI00000CB042; GB:AE006469; PIDN:AAK64939.1;
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMa0537
A;Genome: plasmid

  Query Match             100.0%;  Score 33;  DB 2;  Length 137;
  Best Local Similarity   40.0%;  Pred. No. 3.2e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         85 FATGLLDCGG 94


RESULT 36
G85060
hypothetical protein AT4g04830 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: G85060
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: G85060
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-139 <STO>
A;Cross-references: UNIPROT:Q9ZS91; UNIPARC:UPI00000A7821; GB:NC_001268; NID:g7267241; PI
C;Genetics:
A;Gene: AT4g04830
A;Map position: 4
C;Superfamily: hypothetical protein YCL033c

  Query Match             100.0%;  Score 33;  DB 2;  Length 139;
  Best Local Similarity   40.0%;  Pred. No. 3.3e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         44 FEEGIFDCVG 53


RESULT 37
E85060
hypothetical protein AT4g04810 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: E85060
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: E85060
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-139 <STO>
A;Cross-references: UNIPROT:Q9M0Z5; UNIPARC:UPI00000AB22A; GB:NC_001268; NID:g7267239; PI
C;Genetics:
A;Gene: AT4g04810
A;Map position: 4
C;Superfamily: hypothetical protein YCL033c

  Query Match             100.0%;  Score 33;  DB 2;  Length 139;
  Best Local Similarity   40.0%;  Pred. No. 3.3e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         44 FEEGIFSCIG 53


RESULT 38
T05465
hypothetical protein T8O5.40 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05465
R;Bevan, M.; Wedler, H.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schueller
submitted to the Protein Sequence Database, February 1998
A;Reference number: Z15417
A;Accession: T05465
A;Molecule type: DNA
A;Residues: 1-141 <BEV>
A;Cross-references: UNIPROT:O49706; UNIPARC:UPI00000AAC95; EMBL:AL021890
A;Experimental source: cultivar Columbia; BAC clone T8O5
C;Genetics:
A;Map position: 4
A;Introns: 37/2; 89/3
A;Note: T8O5.40
```

```
C;Superfamily: hypothetical protein YCL033c

  Query Match             100.0%;  Score 33;  DB 2;  Length 141;
  Best Local Similarity   40.0%;  Pred. No. 3.3e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         48 FEEGTYSCAG 57


RESULT 39
H85060
hypothetical protein AT4g04840 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: H85060
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sprin
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: H85060
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-141 <STO>
A;Cross-references: UNIPROT:Q9M0Z4; UNIPARC:UPI00000A6C24; GB:NC_001268; NID:g7267242; P
C;Genetics:
A;Gene: AT4g04840
A;Map position: 4
C;Superfamily: hypothetical protein YCL033c

  Query Match             100.0%;  Score 33;  DB 2;  Length 141;
  Best Local Similarity   40.0%;  Pred. No. 3.3e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         48 FEEGIYCCVG 57


RESULT 40
T26567
hypothetical protein Y26D4A.1 - Caenorhabditis elegans (fragment)
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T26567
R;White, S.
submitted to the EMBL Data Library, September 1999
A;Reference number: Z20234
A;Accession: T26567
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-141 <WIL>
A;Cross-references: UNIPARC:UPI000017BC51; EMBL:AL110478; NID:e1542139; PIDN:CAB54345.1;
A;Experimental source: clone Y26D4A
C;Genetics:
A;Gene: CESP:Y26D4A.1

  Query Match             100.0%;  Score 33;  DB 2;  Length 141;
  Best Local Similarity   40.0%;  Pred. No. 3.3e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db        123 FHYGVFACEG 132


RESULT 41
T36312
hypothetical protein SCE8.13c - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T36312
R;Saunders, D.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, March 1999
A;Reference number: Z21604
A;Accession: T36312
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-142 <SAU>
A;Cross-references: UNIPROT:Q9Z4W4; UNIPARC:UPI00000DAF25; EMBL:AL035654; PIDN:CAB38591.1
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCE8.13c

  Query Match             100.0%;  Score 33;  DB 2;  Length 142;
  Best Local Similarity   40.0%;  Pred. No. 3.3e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         75 FNGGDSGCVG 84


RESULT 42
T05466
hypothetical protein T8O5.50 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05466
R;Bevan, M.; Wedler, H.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schueller
submitted to the Protein Sequence Database, February 1998
A;Reference number: Z15417
A;Accession: T05466
A;Molecule type: DNA
A;Residues: 1-143 <BEV>
A;Cross-references: UNIPROT:O49707; UNIPARC:UPI000009E636; EMBL:AL021890
A;Experimental source: cultivar Columbia; BAC clone T8O5
C;Genetics:
A;Map position: 4
A;Introns: 39/2; 91/3
A;Note: T8O5.50
C;Superfamily: hypothetical protein YCL033c

  Query Match             100.0%;  Score 33;  DB 2;  Length 143;
  Best Local Similarity   40.0%;  Pred. No. 3.3e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         50 FEEGTYSCAG 59


RESULT 43
PC4428
cytochrome P450 4C8 - termite (Mastotermes darwiniensis) (fragment)
C;Species: Mastotermes darwiniensis (termite)
C;Date: 04-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
C;Accession: PC4428
R;Falckh, P.H.J.; Balcombe, W.; Haritos, V.S.; Ahokas, J.T.
Biochem. Biophys. Res. Commun. 241, 579-583, 1997
A;Title: Isolation and identification of a cytochrome P450 sequence in an Australian ter
A;Reference number: PC4428; MUID:98086394; PMID:9425314
A;Accession: PC4428
A;Molecule type: mRNA
A;Residues: 1-144 <FAL>
A;Cross-references: UNIPROT:P91762; UNIPARC:UPI00000773EC; GB:U77126; NID:g1684797; PIDN
C;Comment: This protein is involved in oxidation of a wide range of endogenous and exoge
C;Genetics:
A;Gene: cyp4c8
C;Superfamily: human cytochrome P450 CYP4B1; cytochrome P450 homology
C;Keywords: chromoprotein; heme; iron; metalloprotein
F;140/Binding site: heme iron (Cys) (axial ligand) #status predicted

  Query Match             100.0%;  Score 33;  DB 2;  Length 144;
  Best Local Similarity   40.0%;  Pred. No. 3.4e+02;
```

Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db        133 FSAGPRNCIG 142
```

RESULT 44
S20227
cytochrome P450 2C4 - rabbit (fragment)
N;Alternate names: cytochrome P450PBc4
N;Contains: oxidoreductase (EC 1.-.-.-)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 23-Apr-1993 #sequence_revision 23-Apr-1993 #text_change 09-Jul-2004
C;Accession: S20227; S16715
R;Zhao, J.
submitted to the EMBL Data Library, September 1987
A;Reference number: S20227
A;Accession: S20227
A;Molecule type: mRNA
A;Residues: 1-145 <ZHA1>
A;Cross-references: UNIPROT:Q29507; UNIPARC:UPI0000086985; EMBL:M17026; NID:g164932; PID
R;Zhao, J.; Leighton, J.K.; Kemper, B.
Biochem. Biophys. Res. Commun. 146, 224-231, 1987
A;Title: Characterization of rabbit cytochrome P450IIC4 cDNA and induction by phenobarbi
A;Reference number: S16715; MUID:87270743; PMID:3453118
A;Accession: S16715
A;Molecule type: mRNA
A;Residues: 3-145 <ZHA2>
A;Cross-references: UNIPARC:UPI0000174CD4; EMBL:M17026
C;Genetics:
A;Gene: CYP2C4
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C;Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo
F;90/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 100.0%; Score 33; DB 2; Length 145;
Best Local Similarity 40.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         83 FSAGKRMCVG 92
```

RESULT 45
S36442
reverse transcriptase-like protein - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S36442
R;Szemraj, J.; Plucienniczak, G.; Jaworski, J.; Plucienniczak, A.
submitted to the EMBL Data Library, August 1993
A;Description: Evidence for homological recombination with participation of the bovine a
A;Reference number: S36442
A;Accession: S36442
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-147 <SZE>
A;Cross-references: UNIPROT:O97916; UNIPARC:UPI000017C55A; EMBL:Z25525

Query Match 100.0%; Score 33; DB 2; Length 147;
Best Local Similarity 40.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         25 FEIGKGVCQG 34
```

RESULT 46
G82223
PilB-related protein VC1236 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: G82223
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: G82223
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-148 <HEI>
A;Cross-references: UNIPROT:Q9KSM0; UNIPARC:UPI00000C2F04; GB:AE004203; GB:AE003852; NID
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC1236
A;Map position: 1
C;Superfamily: hypothetical protein YCL033c

Query Match 100.0%; Score 33; DB 2; Length 148;
Best Local Similarity 40.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         58 FEPGLYQCAG 67
```

RESULT 47
S30389
hypothetical protein 2 - human
C;Species: Homo sapiens (man)
C;Date: 07-Oct-1994 #sequence_revision 17-Nov-1995 #text_change 08-Oct-1999
C;Accession: S30389
R;Berger, W.; Meindl, A.; van de Pol, T.J.R.; Cremers, F.P.M.; Ropers, H.H.; Doeerner, C
E.M.; Meitinger, T.
Nature Genet. 1, 199-203, 1992
A;Title: Isolation of a candidate gene for Norrie disease by positional cloning.
A;Reference number: S30388; MUID:93265103; PMID:1303235
A;Accession: S30389
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 1-157 <BER>
A;Cross-references: UNIPARC:UPI000011DF5F; EMBL:X65724; NID:g29946; PIDN:CAA46640.1; PID

Query Match 100.0%; Score 33; DB 2; Length 157;
Best Local Similarity 40.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db        114 FPLGSQCCLG 123
```

RESULT 48
F84769
defender against cell death protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 31-Dec-2004
C;Accession: F84769
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: F84769
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-160 <STO>
A;Cross-references: UNIPARC:UPI0000178A87; GB:AE002093; NID:g3608136; PIDN:AAC36169.1; G
C;Genetics:

```
A;Gene: At2g35520
A;Map position: 2
C;Superfamily: Defender against cell death DAD

  Query Match             100.0%;  Score 33;  DB 2;  Length 160;
  Best Local Similarity   40.0%;  Pred. No. 3.7e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db        104 FLSGVLSCIG 113


RESULT 49
PC4262
cytochrome P450 monooxygenase (EC 1.-.-.-) BF6-2 - fruit fly (Drosophila melanogaster) (
C;Species: Drosophila melanogaster
C;Date: 28-May-1997 #sequence_revision 18-Jul-1997 #text_change 09-Jul-2004
C;Accession: PC4262
R;Maitra, S.; Dombrowski, S.M.; Waters, L.C.; Ganguly, R.
Gene 180, 165-171, 1996
A;Title: Three second chromosome-linked clustered Cyp6 genes show differential constitut
A;Reference number: JC5320; MUID:97128822; PMID:8973362
A;Accession: PC4262
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-162 <MAI>
A;Cross-references: UNIPROT:Q9V769; UNIPARC:UPI000016BC70; GB:L46858; NID:g1478047; PID:
A;Experimental source: strain 91-R
C;Comment: This enzyme is a microsomal enzyme, and is involved in insecticide resistance
C;Keywords: oxidoreductase

  Query Match             100.0%;  Score 33;  DB 2;  Length 162;
  Best Local Similarity   40.0%;  Pred. No. 3.7e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         95 FGDGPRNCIG 104


RESULT 50
I55299
cytochrome P450 PB-1 - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I55299
R;Kimura, H.; Sogawa, K.; Sakai, Y.; Fujii-Kuriyama, Y.
J. Biol. Chem. 264, 2338-2342, 1989
A;Title: Alternative splicing mechanism in a cytochrome P-450 (P-450PB-1) gene generates
A;Reference number: I55299; MUID:89123309; PMID:2914909
A;Accession: I55299
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-169 <RES>
A;Cross-references: UNIPROT:Q64614; UNIPARC:UPI00000E6356; GB:M24239; NID:g341144; PIDN:
C;Genetics:
A;Introns: 62/3; 110/1
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C;Keywords: chromoprotein; heme; iron; metalloprotein; transmembrane protein
F;114/Binding site: heme iron (Cys) (axial ligand) #status predicted

  Query Match             100.0%;  Score 33;  DB 2;  Length 169;
  Best Local Similarity   40.0%;  Pred. No. 3.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db        107 FSAGKRMCAG 116


Search completed: March  8, 2006, 11:24:36
Job time : 41 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2006, 11:17:59 ; Search time 229 Seconds
(without alignments)
30.809 Million cell updates/sec

Title: US-10-751-235-14
Perfect score: 33
Sequence: 1 FXXGXXXCXG 10

Scoring table: BLOSUM62DX:*
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	100.0	18	2	Q13665_HUMAN	Q13665 homo sapien
2	33	100.0	21	2	Q9TX62_HELVI	Q9tx62 heliothis v
3	33	100.0	30	2	Q6V0Y8_MUSSP	Q6v0y8 mus spretus
4	33	100.0	32	1	COAT1_BPIF1	O80295 bacteriopha
5	33	100.0	35	2	Q9P9X1_XYLFA	Q9p9x1 xylella fas
6	33	100.0	36	2	O19047_RABIT	O19047 oryctolagus
7	33	100.0	38	2	Q6LBR9_MOUSE	Q6lbr9 mus musculu
8	33	100.0	40	2	Q9S9I8_ARATH	Q9s9i8 arabidopsis
9	33	100.0	44	2	Q9R1D5_MOUSE	Q9r1d5 mus musculu
10	33	100.0	44	2	Q6LBR7_MOUSE	Q6lbr7 mus musculu
11	33	100.0	45	2	Q14555_HUMAN	Q14555 homo sapien
12	33	100.0	47	2	Q84NZ2_GOSBA	Q84nz2 gossypium b
13	33	100.0	47	2	Q84NZ3_GOSBA	Q84nz3 gossypium b
14	33	100.0	47	2	Q84NZ4_9ROSI	Q84nz4 gossypium r
15	33	100.0	47	2	Q84NZ5_9ROSI	Q84nz5 gossypium h
16	33	100.0	49	2	Q62SA3_BACLD	Q62sa3 bacillus li
17	33	100.0	50	1	VG38_BPMD2	O64229 mycobacteri
18	33	100.0	50	2	Q5K6R0_CRAGI	Q5k6r0 crassostrea
19	33	100.0	50	2	Q70WU5_9LILI	Q70wu5 alstroemeri
20	33	100.0	54	2	Q9YH75_9PERC	Q9yh75 ammodytes t
21	33	100.0	55	2	Q66BR5_YERPS	Q66br5 yersinia ps
22	33	100.0	60	2	Q9YH73_SCOMX	Q9yh73 scophthalmu
23	33	100.0	63	2	Q8GXD1_ARATH	Q8gxd1 arabidopsis
24	33	100.0	64	2	Q58F97_MUSDO	Q58f97 musca domes
25	33	100.0	64	2	Q81160_HPBV0	Q81160 hepatitis b
26	33	100.0	66	2	Q9Y087_GLORO	Q9y087 globodera r
27	33	100.0	66	2	Q9F452_9RHIZ	Q9f452 agrobacteri
28	33	100.0	66	2	Q9I5Y3_PSEAE	Q9i5y3 pseudomonas
29	33	100.0	67	2	Q9TX04_CHITE	Q9tx04 chironomus
30	33	100.0	68	2	Q16868_HUMAN	Q16868 homo sapien
31	33	100.0	68	2	Q75MK4_HUMAN	Q75mk4 homo sapien
32	33	100.0	68	2	Q5DN50_9CAUD	Q5dn50 bacteriopha
33	33	100.0	68	2	Q6GVF0_FRAAN	Q6gvf0 fragaria an
34	33	100.0	68	2	Q76BJ9_HPBV0	Q76bj9 hepatitis b
35	33	100.0	68	2	Q7THR3_HPBV0	Q7thr3 hepatitis b
36	33	100.0	68	2	Q80GT7_HPBV0	Q80gt7 hepatitis b
37	33	100.0	68	2	Q8AZ65_HPBV0	Q8az65 hepatitis b
38	33	100.0	68	2	Q8B4D4_HPBV0	Q8b4d4 hepatitis b
39	33	100.0	68	2	Q991L0_HPBV0	Q99110 hepatitis b
40	33	100.0	68	2	Q9DKQ0_HPBV0	Q9dkq0 hepatitis b
41	33	100.0	68	2	Q4SGY2_TETNG	Q4sgy2 tetraodon n
42	33	100.0	69	2	Q67NC4_SYMTH	Q67nc4 symbiobacte
43	33	100.0	70	2	Q40986_PEA	Q40986 pisum sativ
44	33	100.0	71	2	Q9JWM3_NEIMA	Q9jwm3 neisseria m
45	33	100.0	71	2	Q9JXB9_NEIMB	Q9jxb9 neisseria m
46	33	100.0	71	2	Q8V1R5_9PAPI	Q8v1r5 human papil
47	33	100.0	72	2	Q8IWZ9_HUMAN	Q8iwz9 homo sapien
48	33	100.0	72	2	Q8BC64_9PAPI	Q8bc64 human papil
49	33	100.0	72	2	Q8JSL4_9PAPI	Q8jsl4 chimpanzee
50	33	100.0	72	2	Q9Q2R1_9PAPI	Q9q2r1 human papil
51	33	100.0	73	2	Q96KF0_HUMAN	Q96kf0 homo sapien
52	33	100.0	73	2	Q5ID67_PINTA	Q5id67 pinus taeda
53	33	100.0	73	2	Q5IDC1_PINTA	Q5idc1 pinus taeda
54	33	100.0	73	2	Q8JSM0_9PAPI	Q8jsm0 bovine papi
55	33	100.0	74	2	Q58F98_MUSDO	Q58f98 musca domes
56	33	100.0	74	2	Q4FCQ9_EUSGR	Q4fcq9 eustoma gra
57	33	100.0	74	2	Q84610_CHVP1	Q84610 paramecium
58	33	100.0	74	2	Q9YH76_PLAFE	Q9yh76 platichthys
59	33	100.0	75	2	Q5LI59_BACFN	Q5li59 bacteroides
60	33	100.0	76	2	Q9EPX8_RAT	Q9epx8 rattus norv
61	33	100.0	77	2	Q4TDK7_TETNG	Q4tdk7 tetraodon n
62	33	100.0	78	2	Q7Y2F7_9CAUD	Q7y2f7 bacteriopha
63	33	100.0	79	1	DAD1_MAIZE	O81214 zea mays (m
64	33	100.0	79	2	Q570F4_ARATH	Q570f4 arabidopsis
65	33	100.0	79	2	Q63ND0_BURPS	Q63nd0 burkholderi
66	33	100.0	79	2	Q62AC6_BURMA	Q62ac6 burkholderi
67	33	100.0	80	2	Q38035_BPPHC	Q38035 bacteriopha
68	33	100.0	81	2	Q726M1_DESVH	Q726m1 desulfovibr
69	33	100.0	82	2	Q54X01_DICDI	Q54x01 dictyosteli
70	33	100.0	82	2	Q9TTP6_SHEEP	Q9ttp6 ovis aries
71	33	100.0	83	2	Q5XNT2_ANOGA	Q5xnt2 anopheles g
72	33	100.0	84	2	Q7Z446_HUMAN	Q7z446 homo sapien
73	33	100.0	84	2	P93146_GLYEC	P93146 glycyrrhiza
74	33	100.0	85	2	Q6DQ89_MUSAC	Q6dq89 musa acumin
75	33	100.0	85	2	Q9IA67_BRARE	Q9ia67 brachydanio
76	33	100.0	86	1	GON2_ONCMY	O42241 oncorhynchu
77	33	100.0	86	2	Q5XNS9_ANOGA	Q5xns9 anopheles g
78	33	100.0	86	2	Q5XNU4_ANOGA	Q5xnu4 anopheles g
79	33	100.0	86	2	Q5XNU5_ANOGA	Q5xnu5 anopheles g
80	33	100.0	86	2	Q94FX5_MAIZE	Q94fx5 zea mays (m
81	33	100.0	86	2	Q6XAG0_CORCL	Q6xag0 coregonus c
82	33	100.0	86	2	Q9PT25_ONCMY	Q9pt25 oncorhynchu
83	33	100.0	87	2	Q84552_CHVP1	Q84552 paramecium
84	33	100.0	89	2	Q4L230_PHACH	Q4l230 phanerochae
85	33	100.0	89	2	Q5XNT9_ANOGA	Q5xnt9 anopheles g
86	33	100.0	89	2	Q40985_PEA	Q40985 pisum sativ
87	33	100.0	89	2	Q4RAN1_TETNG	Q4ran1 tetraodon n
88	33	100.0	90	2	Q54EJ7_DICDI	Q54ej7 dictyosteli
89	33	100.0	90	2	Q6GV08_MANSE	Q6gv08 manduca sex
90	33	100.0	90	2	Q4LDW5_TOBAC	Q4ldw5 nicotiana t
91	33	100.0	91	2	Q5XNU9_ANOGA	Q5xnu9 anopheles g
92	33	100.0	91	2	Q5XNV0_ANOGA	Q5xnv0 anopheles g
93	33	100.0	91	2	Q94FX4_MAIZE	Q94fx4 zea mays (m
94	33	100.0	93	2	Q6TXU6_BRARP	Q6txu6 brassica ra
95	33	100.0	94	1	SCYBB_HUMAN	O14625 homo sapien
96	33	100.0	94	2	Q12583_CANMA	Q12583 candida mal
97	33	100.0	94	2	Q53YA3_HUMAN	Q53ya3 homo sapien
98	33	100.0	94	2	Q8MIZ0_MACMU	Q8miz0 macaca mula
99	33	100.0	94	2	Q58D69_BOVIN	Q58d69 bos taurus
100	33	100.0	95	2	Q5XM75_ANOGA	Q5xm75 anopheles g

ALIGNMENTS

```
RESULT 1
Q13665_HUMAN
ID   Q13665_HUMAN PRELIMINARY;      PRT;    18 AA.
AC   Q13665;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   Cardiac myosin binding protein (Fragment).
GN   Name=MyBP-C;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=96083592; PubMed=7493025;
RA   Watkins H., Conner D., Thierfelder L., Jarcho J.A., MacRae C.,
RA   McKenna W.J., Maron B.J., Seidman J.G., Seidman C.E.;
RT   "Mutations in the cardiac myosin binding protein-C gene on chromosome
RT   11 cause familial hypertrophic cardiomyopathy.";
RL   Nat. Genet. 11:434-437(1995).
DR   EMBL; S80805; AAB35661.1; -; mRNA.
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  1891 MW;  3EC842CE85A19C97 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 18;
  Best Local Similarity   40.0%;  Pred. No. 2.3e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FDGGIYVCGG 11


RESULT 2
Q9TX62_HELVI
ID   Q9TX62_HELVI  PRELIMINARY;      PRT;    21 AA.
AC   Q9TX62;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Cytochrome P-450 (Fragment).
OS   Heliothis virescens (Noctuid moth) (Owlet moth).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC   Noctuidae; Heliothinae; Heliothis.
OX   NCBI_TaxID=7102;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=94178422; PubMed=8131898;
RA   Hodgson E., Rose R.L., Goh D.K., Rock G.C., Roe R.M.;
RT   "Insect cytochrome P-450: metabolism and resistance to insecticides.";
RL   Biochem. Soc. Trans. 21:1060-1065(1993).
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
SQ   SEQUENCE   21 AA;  2422 MW;  DF21177CB719B5F2 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 21;
  Best Local Similarity   40.0%;  Pred. No. 2.6e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          8 FCLGPRNCIG 17


RESULT 3
Q6V0Y8_MUSSP
ID   Q6V0Y8_MUSSP PRELIMINARY;      PRT;    30 AA.
AC   Q6V0Y8;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Cytochrome P450 4a10 (Fragment).
GN   Name=Cyp4a10;
OS   Mus spretus (Western wild mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Murinae; Mus.
OX   NCBI_TaxID=10096;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RA   Ruiz-Laguna J., Abril N., Pueyo C.;
RL   Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR   EMBL; AY355164; AAQ55849.1; -; mRNA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   InterPro; IPR002403; EP450IV.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00465; EP450IV.
DR   PRINTS; PR00385; P450.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase.
FT   NON_TER       1      1
FT   NON_TER      30     30
SQ   SEQUENCE   30 AA;  3290 MW;  D3A0937E3E70515E CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 30;
  Best Local Similarity   40.0%;  Pred. No. 3.6e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         11 FSGGARNCIG 20


RESULT 4
COAT1_BPIF1
ID   COAT1_BPIF1    STANDARD;      PRT;    32 AA.
AC   O80295;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Coat protein C, polypeptide I.
GN   Name=VII; Synonyms=7;
OS   Bacteriophage If1.
OC   Viruses; ssDNA viruses; Inoviridae; Inovirus.
OX   NCBI_TaxID=10868;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RA   Hill D.F., Hughes G., McNaughton J.C., Stockwell P.A., Petersen G.B.;
RT   "DNA sequence of the filamentous coliphage If1.";
RL   Submitted (OCT-1993) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBUNIT: Coat protein C is composed of two subunits, polypeptide I
CC       (gene VII) and polypeptide II (gene IX) (By similarity).
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; U02303; AAC62152.1; -; Genomic_DNA.
KW   Capsid protein; Structural protein.
SQ   SEQUENCE   32 AA;  3394 MW;  E0365E0A002B50AA CRC64;

  Query Match             100.0%;  Score 33;  DB 1;  Length 32;
  Best Local Similarity   40.0%;  Pred. No. 3.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         14 FNAGLVICFG 23


RESULT 5
Q9P9X1_XYLFA
ID   Q9P9X1_XYLFA PRELIMINARY;      PRT;    35 AA.
AC   Q9P9X1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=Xf2749;
OS   Xylella fastidiosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xylella.
OX   NCBI_TaxID=2371;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=9a5c;
RX   MEDLINE=20365717; PubMed=10910347; DOI=10.1038/35018003;
RA   Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA   Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA   Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA   Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA   Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA   Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA   Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA   Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA   Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA   Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA   Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA   Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA   Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA   Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA   Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA   Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA   Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA   de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA   Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA   Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA   de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA   da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA   da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA   de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuhako M.H.,
RA   Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA   Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT   "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL   Nature 406:151-159(2000).
DR   EMBL; AE004081; AAF85534.1; -; Genomic_DNA.
DR   PIR; A82520; A82520.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   35 AA;  3787 MW;  D84CD0D69718029E CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 35;
  Best Local Similarity   40.0%;  Pred. No. 4.2e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          5 FDAGAAICVG 14


RESULT 6
O19047_RABIT
ID   O19047_RABIT PRELIMINARY;      PRT;    36 AA.
AC   O19047;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JAN-1998 (TrEMBLrel. 05, Last annotation update)
DE   ORF36.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
OC   Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Stengelin S., Becker W., Maier M., Rosenberger J., Kramer W.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ002005; CAA05134.1; -; Genomic_DNA.
SQ   SEQUENCE   36 AA;  3844 MW;  5BE72DD29C85E956 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 36;
  Best Local Similarity   40.0%;  Pred. No. 4.3e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FRKGLKVCRG 11


RESULT 7
Q6LBR9_MOUSE
ID   Q6LBR9_MOUSE PRELIMINARY;      PRT;    38 AA.
AC   Q6LBR9;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   D3 class I MHC gene (exon 5) (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Domesticus;
RX   PubMed=2584927; DOI=10.1084/jem.170.6.1837;
RA   Brorson K.A., Hunt S.W. III, Hunkapiller T., Sun Y.H., Cheroutre H.,
RA   Nickerson D.A., Hood L.;
RT   "Comparison of exon 5 sequences from 35 class I genes of the BALB/c
RT   mouse.";
RL   J. Exp. Med. 170:1837-1858(1989).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Domesticus;
RA   Brorson K.;
RL   Submitted (AUG-1989) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X16198; CAE82020.1; -; Genomic_DNA.
FT   NON_TER       1      1
FT   NON_TER      38     38
SQ   SEQUENCE   38 AA;  4176 MW;  25428879345A1085 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 38;
  Best Local Similarity   40.0%;  Pred. No. 4.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          7 FQHGNHCCSG 16


RESULT 8
Q9S9I8_ARATH
ID   Q9S9I8_ARATH  PRELIMINARY;      PRT;    40 AA.
AC   Q9S9I8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Modification methylase (EC 2.1.1.73) (Cytosine-specific
DE   methyltransferase) (Fragment).
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
```

```
OC   rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=93281384; PubMed=8389441;
RA   Finnegan E.J., Dennis E.S.;
RT   "Isolation and identification by sequence homology of a putative
RT   cytosine methyltransferase from Arabidopsis thaliana.";
RL   Nucleic Acids Res. 21:2383-2388(1993).
DR   PIR; S35641; S35641.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0006306; P:DNA methylation; IEA.
DR   InterPro; IPR001525; C5_DNA_meth.
DR   Pfam; PF00145; DNA_methylase; 1.
DR   PROSITE; PS00094; C5_MTASE_1; 1.
FT   NON_TER       1      1
FT   NON_CONS     16     17
FT   NON_TER      40     40
SQ   SEQUENCE   40 AA;  4448 MW;  33A7296755138EA7 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 40;
  Best Local Similarity   40.0%;  Pred. No. 4.7e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FINGGPPCQG 11


RESULT 9
Q9R1D5_MOUSE
ID   Q9R1D5_MOUSE PRELIMINARY;      PRT;    44 AA.
AC   Q9R1D5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Phosphodiesterase I/nucleotide pyrophosphatase (Fragment).
GN   Name=Enpp2; Synonyms=Npps2;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=129/SvJ;
RX   MEDLINE=20169177; PubMed=10702660;
RA   Piao J.-H., Matsuda Y., Nakamura H., Sano K.;
RT   "Assignment of Pdnp2, the gene encoding phosphodiesterase I/nucleotide
RT   pyrophosphatase 2, to mouse Chromosome 15D2.";
RL   Cytogenet. Cell Genet. 87:172-174(1999).
DR   EMBL; AF128879; AAD50640.1; -; Genomic_DNA.
DR   MGI; MGI:1321390; Enpp2.
DR   GO; GO:0005615; C:extracellular space; TAS.
FT   NON_TER      44     44
SQ   SEQUENCE   44 AA;  4840 MW;  905EB935D7F1BF1F CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 44;
  Best Local Similarity   40.0%;  Pred. No. 5.2e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         18 FAIGVNLCLG 27


RESULT 10
Q6LBR7_MOUSE
ID   Q6LBR7_MOUSE PRELIMINARY;      PRT;    44 AA.
AC   Q6LBR7;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Q1 class I MHC gene (exon 5) (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Domesticus;
RX   PubMed=2584927; DOI=10.1084/jem.170.6.1837;
RA   Brorson K.A., Hunt S.W. III, Hunkapiller T., Sun Y.H., Cheroutre H.,
RA   Nickerson D.A., Hood L.;
RT   "Comparison of exon 5 sequences from 35 class I genes of the BALB/c
RT   mouse.";
RL   J. Exp. Med. 170:1837-1858(1989).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Domesticus;
RA   Brorson K.;
RL   Submitted (AUG-1989) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X16200; CAE82022.1; -; Genomic_DNA.
FT   NON_TER       1      1
FT   NON_TER      44     44
SQ   SEQUENCE   44 AA;  5026 MW;  13ADBEC0561DA4D6 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 44;
  Best Local Similarity   40.0%;  Pred. No. 5.2e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          7 FQHGNHSCSG 16


RESULT 11
Q14555_HUMAN
ID   Q14555_HUMAN PRELIMINARY;      PRT;    45 AA.
AC   Q14555;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Phosphodiesterase I alpha (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=96163899; PubMed=8586446;
RA   Kawagoe H., Soma O., Goji J., Nishimura N., Narita M., Inazawa J.,
RA   Nakamura H., Sano K.;
RT   "Molecular cloning and chromosomal assignment of the human brain-type
RT   phosphodiesterase I/nucleotide.";
RL   Genomics 30:380-384(1995).
DR   EMBL; D45914; BAA08342.1; -; Genomic_DNA.
FT   NON_TER      45     45
SQ   SEQUENCE   45 AA;  5014 MW;  69C831AD36E5E89D CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 45;
  Best Local Similarity   40.0%;  Pred. No. 5.3e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         18 FAVGVNICLG 27


RESULT 12
Q84NZ2_GOSBA
ID   Q84NZ2_GOSBA PRELIMINARY;      PRT;    47 AA.
AC   Q84NZ2;
```

```
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Putative flavonoid 3'-hydroxylase (Fragment).
OS   Gossypium barbadense (Sea-island cotton) (Egyptian cotton).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC   rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
OX   NCBI_TaxID=3634;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22569057; PubMed=12679546; DOI=10.1093/molbev/msg065;
RA   Senchina D.S., Alvarez I., Cronn R.C., Liu B., Rong J., Noyes R.D.,
RA   Paterson A.H., Wing R.A., Wilkins T.A., Wendel J.F.;
RT   "Rate variation among nuclear genes and the age of polyploidy in
RT   Gossypium.";
RL   Mol. Biol. Evol. 20:633-643(2003).
CC   -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR   EMBL; AF525944; AAP05784.1; -; Genomic_DNA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   InterPro; IPR002401; EP450I.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00463; EP450I.
DR   PROSITE; PS00086; CYTOCHROME_P450; 1.
KW   Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase.
FT   NON_TER       1      1
FT   NON_TER      47     47
SQ   SEQUENCE   47 AA;  5120 MW;  8A53B1DCB6D67951 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 47;
  Best Local Similarity   40.0%;  Pred. No. 5.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         13 FGAGRRICAG 22


RESULT 13
Q84NZ3_GOSBA
ID   Q84NZ3_GOSBA PRELIMINARY;      PRT;    47 AA.
AC   Q84NZ3;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Putative flavonoid 3'-hydroxylase (Fragment).
OS   Gossypium barbadense (Sea-island cotton) (Egyptian cotton).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC   rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
OX   NCBI_TaxID=3634;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22569057; PubMed=12679546; DOI=10.1093/molbev/msg065;
RA   Senchina D.S., Alvarez I., Cronn R.C., Liu B., Rong J., Noyes R.D.,
RA   Paterson A.H., Wing R.A., Wilkins T.A., Wendel J.F.;
RT   "Rate variation among nuclear genes and the age of polyploidy in
RT   Gossypium.";
RL   Mol. Biol. Evol. 20:633-643(2003).
CC   -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR   EMBL; AF525943; AAP05783.1; -; Genomic_DNA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   InterPro; IPR002401; EP450I.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00463; EP450I.
DR   PROSITE; PS00086; CYTOCHROME_P450; 1.
KW   Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase.
FT   NON_TER       1      1
FT   NON_TER      47     47
SQ   SEQUENCE   47 AA;  5106 MW;  8A53B1DCB6D7C951 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 47;
  Best Local Similarity   40.0%;  Pred. No. 5.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         13 FGAGRRICAG 22


RESULT 14
Q84NZ4_9ROSI
ID   Q84NZ4_9ROSI PRELIMINARY;      PRT;    47 AA.
AC   Q84NZ4;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Putative flavonoid 3'-hydroxylase (Fragment).
OS   Gossypium raimondii.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC   rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
OX   NCBI_TaxID=29730;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22569057; PubMed=12679546; DOI=10.1093/molbev/msg065;
RA   Senchina D.S., Alvarez I., Cronn R.C., Liu B., Rong J., Noyes R.D.,
RA   Paterson A.H., Wing R.A., Wilkins T.A., Wendel J.F.;
RT   "Rate variation among nuclear genes and the age of polyploidy in
RT   Gossypium.";
RL   Mol. Biol. Evol. 20:633-643(2003).
CC   -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR   EMBL; AF525942; AAP05782.1; -; Genomic_DNA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   InterPro; IPR002401; EP450I.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00463; EP450I.
DR   PROSITE; PS00086; CYTOCHROME_P450; 1.
KW   Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase.
FT   NON_TER       1      1
FT   NON_TER      47     47
SQ   SEQUENCE   47 AA;  5120 MW;  8A53B1DCB6D67951 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 47;
  Best Local Similarity   40.0%;  Pred. No. 5.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         13 FGAGRRICAG 22


RESULT 15
Q84NZ5_9ROSI
ID   Q84NZ5_9ROSI PRELIMINARY;      PRT;    47 AA.
AC   Q84NZ5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Putative flavonoid 3'-hydroxylase (Fragment).
OS   Gossypium herbaceum.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC   rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
OX   NCBI_TaxID=34274;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
```

```
RX   MEDLINE=22569057; PubMed=12679546; DOI=10.1093/molbev/msg065;
RA   Senchina D.S., Alvarez I., Cronn R.C., Liu B., Rong J., Noyes R.D.,
RA   Paterson A.H., Wing R.A., Wilkins T.A., Wendel J.F.;
RT   "Rate variation among nuclear genes and the age of polyploidy in
RT   Gossypium.";
RL   Mol. Biol. Evol. 20:633-643(2003).
CC   -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR   EMBL; AF525941; AAP05781.1; -; Genomic_DNA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   InterPro; IPR002401; EP450I.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00463; EP450I.
DR   PROSITE; PS00086; CYTOCHROME_P450; 1.
KW   Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase.
FT   NON_TER       1      1
FT   NON_TER      47     47
SQ   SEQUENCE   47 AA;  5106 MW;  8A53B1DCB6D7C951 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 47;
  Best Local Similarity   40.0%;  Pred. No. 5.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         13 FGAGRRICAG 22


RESULT 16
Q62SA3_BACLD
ID   Q62SA3_BACLD PRELIMINARY;      PRT;    49 AA.
AC   Q62SA3;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=BL02399;
OS   Bacillus licheniformis (strain DSM 13 / ATCC 14580).
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=279010;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15461803;
RA   Rey M.W., Ramaiya P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
RA   Tang M., Lopez de Leon A., Xiang H., Gusti V., Clausen I.G.,
RA   Olsen P.B., Rasmussen M.D., Andersen J.T., Joergensen P.L.,
RA   Larsen T.S., Sorokin A., Bolotin A., Lapidus A., Galleron N.,
RA   Ehrlich S.D., Berka R.M.;
RT   "Complete genome sequence of the industrial bacterium Bacillus
RT   licheniformis and comparisons with closely related Bacillus species.";
RL   Genome Biol. 5:RESEARCH077.1-RESEARCH077.12(2004).
DR   EMBL; CP000002; AAU24356.1; -; Genomic_DNA.
DR   InterPro; IPR010070; Cons_hypoth_TTM.
DR   TIGRFAMs; TIGR01732; tiny_TM_bacill; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   49 AA;  4794 MW;  5FDF1724D0001D5B CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 49;
  Best Local Similarity   40.0%;  Pred. No. 5.7e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          3 FGYGGGCCGG 12


RESULT 17
VG38_BPMD2
ID   VG38_BPMD2     STANDARD;      PRT;    50 AA.
AC   O64229;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Gene 38 protein (Gp38).
GN   Name=38;
OS   Mycobacteriophage D29.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC   L5-like viruses.
OX   NCBI_TaxID=28369;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   MEDLINE=98300335; PubMed=9636706; DOI=10.1006/jmbi.1997.1610;
RA   Ford M.E., Sarkis G.J., Belanger A.E., Hendrix R.W., Hatfull G.F.;
RT   "Genome structure of mycobacteriophage D29: implications for phage
RT   evolution.";
RL   J. Mol. Biol. 279:143-164(1998).
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; AF022214; AAC18479.1; -; Genomic_DNA.
DR   PIR; D72804; D72804.
SQ   SEQUENCE   50 AA;  4851 MW;  75BCC1A1CF2EF26E CRC64;

  Query Match             100.0%;  Score 33;  DB 1;  Length 50;
  Best Local Similarity   40.0%;  Pred. No. 5.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         13 FALGLTACDG 22


RESULT 18
Q5K6R0_CRAGI
ID   Q5K6R0_CRAGI PRELIMINARY;      PRT;    50 AA.
AC   Q5K6R0;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Cytochrome P450 related protein (Fragment).
OS   Crassostrea gigas (Pacific oyster).
OC   Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia; Ostreoida;
OC   Ostreoidea; Ostreidae; Crassostrea.
OX   NCBI_TaxID=29159;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Cardiac;
RA   Lopes E., Ohresser M.C.P., Cancela M.L.;
RT   "Identification and expected function of some genes differentially
RT   expressed in either heart or heart-derived cell cultures in
RT   Crassostrea gigas: a new approach to better understanding Bivalvia
RT   primary cell culture.";
RL   Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR   EMBL; AF075692; AAQ13473.1; -; mRNA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   InterPro; IPR002401; EP450I.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00463; EP450I.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase.
FT   NON_TER       1      1
FT   NON_TER      50     50
SQ   SEQUENCE   50 AA;  5802 MW;  191409A96CDC85BA CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 50;
  Best Local Similarity   40.0%;  Pred. No. 5.8e+02;
```

```
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         31 FGVGRRRCVG 40


RESULT 19
Q70WU5_9LILI
ID   Q70WU5_9LILI PRELIMINARY;      PRT;    50 AA.
AC   Q70WU5;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Dad-1 homologue (Fragment).
GN   Name=dad1;
OS   Alstroemeria hybrid cultivar 'Samora'.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Alstroemeriaceae;
OC   Alstroemeria.
OX   NCBI_TaxID=160840;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Petal;
RA   Wagstaff C., Malcolm P., Rafiq A., Leverentz M., Griffiths G.,
RA   Thomas B., Stead A., Rogers H.;
RT   "Programmed cell death (PCD) processes begin extremely early in
RT   Alstroemeria petal senescence.";
RL   New Phytol. 160:49-59(2003).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Petal;
RA   Rogers H.J.;
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ514409; CAD55820.1; -; mRNA.
DR   InterPro; IPR003038; DAD.
DR   Pfam; PF02109; DAD; 1.
FT   NON_TER       1      1
FT   NON_TER      50     50
SQ   SEQUENCE   50 AA;  5463 MW;  46455530700E0A80 CRC64;

  Query Match           100.0%;  Score 33;  DB 2;  Length 50;
  Best Local Similarity   40.0%;  Pred. No. 5.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          8 FLSGVLSCVG 17


RESULT 20
Q9YH75_9PERC
ID   Q9YH75_9PERC PRELIMINARY;      PRT;    54 AA.
AC   Q9YH75;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Cytochrome P450 1A (Fragment).
GN   Name=CYP1A1;
OS   Ammodytes tobianus.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;
OC   Trachinoidei; Ammodytidae; Ammodytes.
OX   NCBI_TaxID=84621;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RA   Robertson F.E., McPhail M.E., Stagg R.M., Craft J.A.;
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR   EMBL; AJ130768; CAA10203.1; -; mRNA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   InterPro; IPR002401; EP450I.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00463; EP450I.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase.
FT   NON_TER       1      1
FT   NON_TER      54     54
SQ   SEQUENCE   54 AA;  6181 MW;  0429CF02A404D8E8 CRC64;

  Query Match           100.0%;  Score 33;  DB 2;  Length 54;
  Best Local Similarity   40.0%;  Pred. No. 6.3e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         35 FGLGKRRCIG 44


RESULT 21
Q66BR5_YERPS
ID   Q66BR5_YERPS PRELIMINARY;      PRT;    55 AA.
AC   Q66BR5;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Putative phage minor tail protein.
GN   OrderedLocusNames=YPTB1705;
OS   Yersinia pseudotuberculosis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Yersinia.
OX   NCBI_TaxID=633;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=IP32953 / Serotype I;
RX   PubMed=15358858; DOI=10.1073/pnas.0404012101;
RA   Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA   Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA   Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
RA   Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M.,
RA   Derbise A., Hauser L.J., Garcia E.;
RT   "Insights into the evolution of Yersinia pestis through whole-genome
RT   comparison with Yersinia pseudotuberculosis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR   EMBL; BX936398; CAH20944.1; -; Genomic_DNA.
KW   Complete proteome.
SQ   SEQUENCE   55 AA;  6624 MW;  D20E1ADD3FF2F29D CRC64;

  Query Match           100.0%;  Score 33;  DB 2;  Length 55;
  Best Local Similarity   40.0%;  Pred. No. 6.4e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         24 FVLGHTHCWG 33


RESULT 22
Q9YH73_SCOMX
ID   Q9YH73_SCOMX PRELIMINARY;      PRT;    60 AA.
AC   Q9YH73;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Cytochrome P450 1A (Fragment).
GN   Name=CYP1A1;
OS   Scophthalmus maximus (Turbot).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC   Pleuronectoidei; Scophthalmidae; Scophthalmus.
```

```
OX   NCBI_TaxID=52904;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RA   Robertson F.E., McPhail M.E., Stagg R.M., Craft J.A.;
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR   EMBL; AJ130770; CAA10205.1; -; mRNA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   InterPro; IPR002403; EP450IV.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00465; EP450IV.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Heme; Iron; Metal-binding.
FT   NON_TER       1      1
FT   NON_TER      60     60
SQ   SEQUENCE   60 AA;  6988 MW;  5379CF03ED02A142 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 60;
  Best Local Similarity   40.0%;  Pred. No. 6.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         41 FGLGKRRCIG 50


RESULT 23
Q8GXD1_ARATH
ID   Q8GXD1_ARATH PRELIMINARY;      PRT;    63 AA.
AC   Q8GXD1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein At1g57750/T8L23_21.
GN   Name=At1g57750/T8L23_21;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC   rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
RA   Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA   Hayashizaki Y., Shinozaki K.;
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK118299; BAC42917.1; -; mRNA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   Pfam; PF00067; p450; 1.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Hypothetical protein.
SQ   SEQUENCE   63 AA;  7063 MW;  5D00C1D972039F85 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 63;
  Best Local Similarity   40.0%;  Pred. No. 7.2e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          3 FNSGPRTCLG 12


RESULT 24
Q58F97_MUSDO
ID   Q58F97_MUSDO PRELIMINARY;      PRT;    64 AA.
AC   Q58F97;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Cytochrome P450 (Fragment).
OS   Musca domestica (House fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Muscoidea;
OC   Muscidae; Musca.
OX   NCBI_TaxID=7370;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Ma C.X., Qiu X.H., Li M., He F.Q., Liu H.X.;
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR   EMBL; AY960134; AAX54685.1; -; mRNA.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase.
FT   NON_TER       1      1
SQ   SEQUENCE   64 AA;  7164 MW;  597E70ECAF195DF8 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 64;
  Best Local Similarity   40.0%;  Pred. No. 7.3e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FGEGPRNCIG 11


RESULT 25
Q81160_HPBV0
ID   Q81160_HPBV0 PRELIMINARY;      PRT;    64 AA.
AC   Q81160;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Surface antigen (HBsAg).
OS   Hepatitis B virus.
OC   Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC   Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=85107103; PubMed=3968537;
RA   Takeshima H., Inokoshi J., Namiki M., Shimada J., Omura S.;
RT   "Structural analysis of the gene coding for hepatitis B virus surface
RT   antigen and its product.";
RL   J. Gen. Virol. 66:195-200(1985).
DR   EMBL; M23808; AAA45500.1; -; mRNA.
DR   GO; GO:0016032; P:viral life cycle; IEA.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   PANTHER; PTHR10832; Hepvir_surfAg; 1.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
SQ   SEQUENCE   64 AA;  6842 MW;  122E2B1D0404D755 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 64;
  Best Local Similarity   40.0%;  Pred. No. 7.3e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         41 FLGGAPTCLG 50


RESULT 26
Q9Y087_GLORO
ID   Q9Y087_GLORO PRELIMINARY;      PRT;    66 AA.
AC   Q9Y087;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
```

```
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Aldehyde dehydrogenase (Fragment).
GN   Name=adh-1;
OS   Globodera rostochiensis (Golden nematode worm).
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
OC   Tylenchoidea; Heteroderidae; Heteroderinae; Globodera.
OX   NCBI_TaxID=31243;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Qin L., Schots A., Smant G., Bakker J., Helder J.;
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ133109; CAB45527.1; -; mRNA.
DR   HSSP; P05091; 1O02.
DR   SMR; Q9Y087; 10-66.
DR   GO; GO:0016491; F:oxidoreductase activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR002086; Aldehyd_dehydrog.
DR   Pfam; PF00171; Aldedh; 1.
DR   PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR   PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
KW   Oxidoreductase.
FT   NON_TER       1      1
FT   NON_TER      66     66
SQ   SEQUENCE   66 AA;  7168 MW;  FAAFAF2FA8F831A8 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 66;
  Best Local Similarity   40.0%;  Pred. No. 7.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         45 FNQGQCCCAG 54


RESULT 27
Q9F452_9RHIZ
ID   Q9F452_9RHIZ PRELIMINARY;      PRT;    66 AA.
AC   Q9F452;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   P-450-type monooxygenase (Fragment).
GN   Name=virH2;
OS   Agrobacterium tumefaciens.
OG   Plasmid pTi15955.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX   NCBI_TaxID=358;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=88247765; PubMed=2837739;
RA   Thompson D.V., Melchers L.S., Idler K.B., Shilperoort R.A.,
RA   Hooykaas P.J.J.;
RT   "Analysis of the complete nucleotide sequence of the Agrobacterium
RT   tumefaciens virB operon.";
RL   Nucleic Acids Res. 16:4621-4636(1988).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Schrammeijer B., Beijersbergen A., Idler K.B., Melchers L.S.,
RA   Thompson D.V., Hooykaas P.J.J.;
RT   "Sequence analysis of the vir-region from Agrobacterium tumefaciens
RT   octopine Ti plasmid pTi15955.";
RL   J. Exp. Bot. 347:1167-1169(2000).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RA   Thompson D.;
RL   Submitted (NOV-1990) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RA   Schrammeijer B.;
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR   EMBL; X06826; CAC15160.1; -; Genomic_DNA.
DR   HSSP; Q9L142; 1GWI.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase; Plasmid.
FT   NON_TER       1      1
SQ   SEQUENCE   66 AA;  7329 MW;  824AAC9FE080AADF CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 66;
  Best Local Similarity   40.0%;  Pred. No. 7.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          8 FGSGPHHCPG 17


RESULT 28
Q9I5Y3_PSEAE
ID   Q9I5Y3_PSEAE PRELIMINARY;      PRT;    66 AA.
AC   Q9I5Y3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=PA0553;
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H. Jr., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
DR   EMBL; AE004492; AAG03942.1; -; Genomic_DNA.
DR   PIR; A83575; A83575.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   66 AA;  7030 MW;  67FFE399CDF97712 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 66;
  Best Local Similarity   40.0%;  Pred. No. 7.5e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         11 FCLGLAACSG 20


RESULT 29
Q9TX04_CHITE
ID   Q9TX04_CHITE PRELIMINARY;      PRT;    67 AA.
AC   Q9TX04;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   CE75=E75A ecdysteroid receptor homolog (Fragment).
OS   Chironomus tentans (Midge).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Nematocera; Chironomoidea;
OC   Chironomidae; Chironominae; Chironomus.
OX   NCBI_TaxID=7153;
RN   [1]
```

```
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=96088569; PubMed=7579577;
RA   Wegmann I.S., Quack S., Spindler K.D., Dorsch-Hasler K., Vogtli M.,
RA   Lezzi M.;
RT   "Immunological studies on the developmental and chromosomal
RT   distribution of ecdysteroid receptor protein in Chironomus tentans.";
RL   Arch. Insect Biochem. Physiol. 30:95-114(1995).
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- SIMILARITY: Belongs to the nuclear hormone receptor family.
DR   HSSP; P03372; 1HCQ.
DR   SMR; Q9TX04; 1-67.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004879; F:ligand-dependent nuclear receptor activity; IEA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   InterPro; IPR000324; VitD_receptor.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00047; STROIDFINGER.
DR   PRINTS; PR00350; VITAMINDR.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_REC_DBD_1; 1.
KW   DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW   Transcription regulation; Zinc; Zinc-finger.
FT   NON_TER       1      1
FT   NON_TER      67     67
SQ   SEQUENCE   67 AA;  7822 MW;  42558F74AD1AE611 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 67;
  Best Local Similarity   40.0%;  Pred. No. 7.6e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         11 FHYGVHSCEG 20


RESULT 30
Q16868_HUMAN
ID   Q16868_HUMAN PRELIMINARY;      PRT;    68 AA.
AC   Q16868;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   CYP2E1 protein (Fragment).
GN   Name=CYP2E1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=95251674; PubMed=7733943;
RA   Seree E.M., Botto F., Pisano P., Lechevalier E., Desobry A., Barra Y.;
RT   "Evidence for the existence of two human CYP2E1 cDNAs using different
RT   polyadenylation signals.";
RL   Biochem. Biophys. Res. Commun. 209:717-722(1995).
CC   -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR   EMBL; S77873; AAD14267.1; -; mRNA.
DR   HSSP; P11712; 1OG2.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   InterPro; IPR002401; EP450I.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00463; EP450I.
DR   PROSITE; PS00086; CYTOCHROME_P450; 1.
KW   Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase.
FT   NON_TER      68     68
SQ   SEQUENCE   68 AA;  7602 MW;  594C69404F84E0F2 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 68;
  Best Local Similarity   40.0%;  Pred. No. 7.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          5 FSAGKRVCAG 14


RESULT 31
Q75MK4_HUMAN
ID   Q75MK4_HUMAN PRELIMINARY;      PRT;    68 AA.
AC   Q75MK4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein STK17A (Fragment).
GN   Name=STK17A;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22737999; PubMed=12853948; DOI=10.1038/nature01782;
RA   Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,
RA   Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R.,
RA   Wylie K., Sekhon M., Becker M.C., O'Laughlin M.D., Schaller M.E.,
RA   Fewell G.A., Delehaunty K.D., Miner T.L., Nash W.E., Cordes M., Du H.,
RA   Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Andrews S., Isak A.,
RA   Vanbrunt A., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J.,
RA   Ozersky P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A.,
RA   Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S.,
RA   Kozlowicz-Reilly A., Leonard S., Rohlfing T., Rock S.M.,
RA   Tin-Wollam A.-M., Abbott A., Minx P., Maupin R., Strowmatt C.,
RA   Latreille P., Miller N., Johnson D., Murray J., Woessner J.P.,
RA   Wendl M.C., Yang S.-P., Schultz B.R., Wallis J.W., Spieth J.,
RA   Bieri T.A., Nelson J.O., Berkowicz N., Wohldmann P.E., Cook L.L.,
RA   Hickenbotham M.T., Eldred J., Williams D., Bedell J.A., Mardis E.R.,
RA   Clifton S.W., Chissoe S.L., Marra M.A., Raymond C., Haugen E.,
RA   Gillett W., Zhou Y., James R., Phelps K., Iadanoto S., Bubb K.,
RA   Simms E., Levy R., Clendenning J., Kaul R., Kent W.J., Furey T.S.,
RA   Baertsch R.A., Brent M.R., Keibler E., Flicek P., Bork P., Suyama M.,
RA   Bailey J.A., Portnoy M.E., Torrents D., Chinwalla A.T., Gish W.R.,
RA   Eddy S.R., McPherson J.D., Olson M.V., Eichler E.E., Green E.D.,
RA   Waterston R.H., Wilson R.K.;
RT   "The DNA sequence of human chromosome 7.";
RL   Nature 424:157-164(2003).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Glaser E., Lamar B., LaPlant Y., Stoneking T.;
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RA   Waterston R.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RA   Wilson R.;
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC011738; AAS02010.1; -; Genomic_DNA.
KW   Hypothetical protein.
FT   NON_TER      68     68
SQ   SEQUENCE   68 AA;  6886 MW;  D298AB1D7BBEC5A8 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 68;
  Best Local Similarity   40.0%;  Pred. No. 7.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
```

```
Db         55 FQDGYSLCPG 64


RESULT 32
Q5DN50_9CAUD
ID   Q5DN50_9CAUD PRELIMINARY;      PRT;    68 AA.
AC   Q5DN50;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Gp55.
GN   ORFNames=JL001p55;
OS   Bacteriophage phi JL001.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX   NCBI_TaxID=279383;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Lohr J.E., Chen F., Hill R.T.;
RT   "Genomic Analysis of Bacteriophage PhiJL001: Insights into Its
RT   Interaction with a Sponge-Associated Alpha-Proteobacterium.";
RL   Appl. Environ. Microbiol. 71:1598-1609(2005).
DR   EMBL; AY576273; AAT69531.1; -; Genomic_DNA.
SQ   SEQUENCE   68 AA;  7666 MW;  78D629C164309F70 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 68;
  Best Local Similarity   40.0%;  Pred. No. 7.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         26 FGGGWEFCLG 35


RESULT 33
Q6GVF0_FRAAN
ID   Q6GVF0_FRAAN PRELIMINARY;      PRT;    68 AA.
AC   Q6GVF0;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Cytochrome P450 (Fragment).
OS   Fragaria ananassa (Strawberry).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC   rosids; eurosids I; Rosales; Rosaceae; Rosoideae; Fragaria.
OX   NCBI_TaxID=3747;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Achene;
RA   Balogh A., Koncz T., Kiss E., Heszky L.E.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR   EMBL; AY633994; AAT46620.1; -; mRNA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   InterPro; IPR002401; EP450I.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00463; EP450I.
DR   PROSITE; PS00086; CYTOCHROME_P450; 1.
KW   Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase.
FT   NON_TER       1      1
FT   NON_TER      68     68
SQ   SEQUENCE   68 AA;  7416 MW;  9D232752B90EB217 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 68;
  Best Local Similarity   40.0%;  Pred. No. 7.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          6 FGAGRRICPG 15


RESULT 34
Q76BJ9_HPBV0
ID   Q76BJ9_HPBV0 PRELIMINARY;      PRT;    68 AA.
AC   Q76BJ9;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   HBs antigen.
OS   Hepatitis B virus.
OC   Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC   Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Tanaka Y., Yeo A.E., Orito E., Ito K., Hirashima N., Ide T., Sata M.,
RA   Mizokami M.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB111113; BAD02901.1; -; Genomic_DNA.
DR   GO; GO:0016032; P:viral life cycle; IEA.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   PANTHER; PTHR10832; Hepvir_surfAg; 1.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
SQ   SEQUENCE   68 AA;  7163 MW;  7351D6EAA55E17AD CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 68;
  Best Local Similarity   40.0%;  Pred. No. 7.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         41 FLGGAPACPG 50


RESULT 35
Q7THR3_HPBV0
ID   Q7THR3_HPBV0 PRELIMINARY;      PRT;    68 AA.
AC   Q7THR3;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Truncated S protein.
OS   Hepatitis B virus.
OC   Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC   Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=MH534978;
RA   Zhang J.M., Wen Y.M.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY220700; AAP40757.1; -; Genomic_DNA.
DR   GO; GO:0016032; P:viral life cycle; IEA.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   PANTHER; PTHR10832; Hepvir_surfAg; 1.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
SQ   SEQUENCE   68 AA;  7220 MW;  668BE6EAA545161D CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 68;
  Best Local Similarity   40.0%;  Pred. No. 7.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         41 FLGGAPTCPG 50


RESULT 36
Q80GT7_HPBV0
ID   Q80GT7_HPBV0 PRELIMINARY;      PRT;    68 AA.
```

```
AC   Q80GT7;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Truncated S protein.
OS   Hepatitis B virus.
OC   Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC   Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=FMU022;
RA   Lin X., Zheng D.L., Xu X.;
RL   Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY206393; AAP06652.1; -; Genomic_DNA.
DR   GO; GO:0016032; P:viral life cycle; IEA.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   PANTHER; PTHR10832; Hepvir_surfAg; 1.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
SQ   SEQUENCE   68 AA;  7235 MW;  777BE6EAB994DB1D CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 68;
  Best Local Similarity   40.0%;  Pred. No. 7.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         41 FLGGAPTCPG 50


RESULT 37
Q8AZ65_HPBV0
ID   Q8AZ65_HPBV0 PRELIMINARY;      PRT;    68 AA.
AC   Q8AZ65;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Clone S6.5, complete genome (Clone S6.7, complete genome).
OS   Hepatitis B virus.
OC   Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC   Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Vaishali C., Acharya S.K., Panda S.K.;
RT   "Cryptic Hepatitis B virus infection: analysis of the complete genomic
RT   sequence of HBV from nine patients with seronegative viral
RT   hepatitis.";
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY161154; AAO12677.1; -; Genomic_DNA.
DR   EMBL; AY161156; AAO12688.1; -; Genomic_DNA.
DR   GO; GO:0016032; P:viral life cycle; IEA.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   PANTHER; PTHR10832; Hepvir_surfAg; 1.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
SQ   SEQUENCE   68 AA;  7323 MW;  10DF16EF5B6C073F CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 68;
  Best Local Similarity   40.0%;  Pred. No. 7.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         41 FLGGTTVCLG 50


RESULT 38
Q8B4D4_HPBV0
ID   Q8B4D4_HPBV0 PRELIMINARY;      PRT;    68 AA.
AC   Q8B4D4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Clone S6.2, complete genome.
OS   Hepatitis B virus.
OC   Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC   Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Vaishali C., Acharya S.K., Panda S.K.;
RT   "Cryptic Hepatitis B virus infection: analysis of the complete genomic
RT   sequence of HBV from nine patients with seronegative viral
RT   hepatitis.";
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY161151; AAO12668.1; -; Genomic_DNA.
DR   GO; GO:0016032; P:viral life cycle; IEA.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   PANTHER; PTHR10832; Hepvir_surfAg; 1.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen.
SQ   SEQUENCE   68 AA;  7295 MW;  D3DF16EF59276A8B CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 68;
  Best Local Similarity   40.0%;  Pred. No. 7.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         41 FLGGTTVCLG 50


RESULT 39
Q991L0_HPBV0
ID   Q991L0_HPBV0 PRELIMINARY;      PRT;    68 AA.
AC   Q991L0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Truncated envelope protein.
OS   Hepatitis B virus.
OC   Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC   Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B11;
RA   Dong J., Cheng J., Wang Q.H., Shi S.S., Hong Y., Huangfu J.K.,
RA   Wang G., Li L., Si C.W.;
RT   "The study on quasispecies of hepatitis b virus: reverse transcriptase
RT   region in polymerase gene as an example.";
RL   Ping Tu Hsueh Pao 17:270-272(2001).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B11;
RA   Dong J., Cheng J., Huangfu J.K., Hong Y., Wang G., Chen C.G., Li L.,
RA   Zhang L.X., Chen J.M.;
RT   "The preliminary study on individually characterized quasispecies of
RT   hepatitis B virus.";
RL   Jie Fang Jun Yi Xue Za Zhi 27:119-121(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B11;
RA   Dong J., Cheng J., Wang Q.H., Shi S.S., Hong Y., Huangfu J.K.,
RA   Wang G., Li L., Si C.W.;
RT   "The study on quasispecies of hepatitis B virus: reverse transcriptase
RT   region in polymerase gene as an example.";
RL   Jie Fang Jun Yi Xue Za Zhi 26:823-825(2002).
DR   EMBL; AF335734; AAK19537.1; -; Genomic_DNA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0016032; P:viral life cycle; IEA.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   PANTHER; PTHR10832; Hepvir_surfAg; 1.
DR   Pfam; PF00695; vMSA; 1.
```

```
KW   Antigen; Envelope protein.
SQ   SEQUENCE   68 AA;  7182 MW;  A59A071AB6BA159D CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 68;
  Best Local Similarity   40.0%;  Pred. No. 7.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         41 FLGGASTCPG 50


RESULT 40
Q9DKQ0_HPBV0
ID   Q9DKQ0_HPBV0 PRELIMINARY;      PRT;    68 AA.
AC   Q9DKQ0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Mutant envelope protein.
OS   Hepatitis B virus.
OC   Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC   Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Dong J., Cheng J., Wang Q.H., Shi S.S., Hong Y., Huangfu J.K.,
RA   Wang G., Li L., Si C.W.;
RT   "The study on quasispecies of hepatitis b virus: reverse transcriptase
RT   region in polymerase gene as an example.";
RL   Ping Tu Hsueh Pao 17:270-272(2001).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Dong J., Cheng J., Huangfu J.K., Hong Y., Wang G., Chen C.G., Li L.,
RA   Zhang L.X., Chen J.M.;
RT   "The preliminary study on individually characterized quasispecies of
RT   hepatitis B virus.";
RL   Jie Fang Jun Yi Xue Za Zhi 27:119-121(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RA   Dong J., Cheng J., Wang Q.H., Shi S.S., Hong Y., Huangfu J.K.,
RA   Wang G., Li L., Si C.W.;
RT   "The study on quasispecies of hepatitis B virus: reverse transcriptase
RT   region in polymerase gene as an example.";
RL   Jie Fang Jun Yi Xue Za Zhi 26:823-825(2002).
DR   EMBL; AF329859; AAG48739.1; -; Genomic_DNA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0016032; P:viral life cycle; IEA.
DR   InterPro; IPR000349; Hepvir_surfAg.
DR   PANTHER; PTHR10832; Hepvir_surfAg; 1.
DR   Pfam; PF00695; vMSA; 1.
KW   Antigen; Envelope protein.
SQ   SEQUENCE   68 AA;  7220 MW;  669A071AA8752B1D CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 68;
  Best Local Similarity   40.0%;  Pred. No. 7.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         41 FLGGAPTCPG 50


RESULT 41
Q4SGY2_TETNG
ID   Q4SGY2_TETNG PRELIMINARY;      PRT;    68 AA.
AC   Q4SGY2;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Chromosome 14 SCAF14590, whole genome shotgun sequence.
GN   ORFNames=GSTENG00018412001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAE01014590; CAG00100.1; -; Genomic_DNA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   InterPro; IPR002401; EP450I.
DR   PRINTS; PR00463; EP450I.
KW   Heme; Iron; Membrane; Metal-binding; Monooxygenase; Oxidoreductase.
SQ   SEQUENCE   68 AA;  7920 MW;  E367B1591080F9B6 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 68;
  Best Local Similarity   40.0%;  Pred. No. 7.8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         30 FALGPRSCLG 39


RESULT 42
Q67NC4_SYMTH
ID   Q67NC4_SYMTH PRELIMINARY;      PRT;    69 AA.
AC   Q67NC4;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=STH1834;
OS   Symbiobacterium thermophilum.
OC   Bacteria; Actinobacteria; Symbiobacterium.
OX   NCBI_TaxID=2734;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=IAM14863;
RX   PubMed=15383646; DOI=10.1093/nar/gkh830;
RA   Ueda K., Yamashita A., Ishikawa J., Shimada M., Watsuji T.,
RA   Morimura K., Ikeda H., Hattori M., Beppu T.;
RT   "Genome sequence of Symbiobacterium thermophilum, an uncultivable
RT   bacterium that depends on microbial commensalism.";
RL   Nucleic Acids Res. 32:4937-4944(2004).
DR   EMBL; AP006840; BAD40819.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   69 AA;  7618 MW;  50F6B384E1DF1C3F CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 69;
  Best Local Similarity   40.0%;  Pred. No. 7.9e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         49 FALGDASCSG 58


RESULT 43
Q40986_PEA
ID   Q40986_PEA  PRELIMINARY;      PRT;    70 AA.
AC   Q40986;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Pisum sativum clone MFRTPCR62 wound-inducible cytochrome P450
DE   (Fragment).
OS   Pisum sativum (Garden pea).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae; Pisum.
OX   NCBI_TaxID=3888;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Wounded;
RX   MEDLINE=96417083; PubMed=8819874; DOI=10.1104/pp.110.3.1035;
RA   Frank M.R., Deyneka J.M., Schuler M.A.;
RT   "Cloning of wound-induced cytochrome P450 monooxygenases expressed in
RT   pea.";
RL   Plant Physiol. 110:1035-1046(1996).
DR   EMBL; U29335; AAC49190.1; -; mRNA.
DR   PIR; T06525; T06525.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
FT   NON_TER       1      1
SQ   SEQUENCE   70 AA;  7952 MW;  F8E8E7CA1752CA23 CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 70;
  Best Local Similarity   40.0%;  Pred. No. 8e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          6 FQAGPRVCLG 15


RESULT 44
Q9JWM3_NEIMA
ID   Q9JWM3_NEIMA PRELIMINARY;      PRT;    71 AA.
AC   Q9JWM3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative integral membrane protein.
GN   OrderedLocusNames=NMA0300;
OS   Neisseria meningitidis (serogroup A).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=65699;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX   MEDLINE=20222556; PubMed=10761919; DOI=10.1038/35006655;
RA   Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,
RA   Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA   Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA   Jagels K., Leather S., Moule S., Mungall K.L., Quail M.A.,
RA   Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA   Whitehead S., Spratt B.G., Barrell B.G.;
RT   "Complete DNA sequence of a serogroup A strain of Neisseria
RT   meningitidis Z2491.";
RL   Nature 404:502-506(2000).
DR   EMBL; AL162752; CAB83605.1; -; Genomic_DNA.
DR   PIR; D82025; D82025.
KW   Complete proteome.
SQ   SEQUENCE   71 AA;  8104 MW;  50EE1A62D021F67E CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 71;
  Best Local Similarity   40.0%;  Pred. No. 8.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FKLGVYACLG 11


RESULT 45
Q9JXB9_NEIMB
ID   Q9JXB9_NEIMB PRELIMINARY;      PRT;    71 AA.
AC   Q9JXB9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=NMB2131;
OS   Neisseria meningitidis (serogroup B).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=491;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=MC58 / Serogroup B;
RX   MEDLINE=20175755; PubMed=10710307; DOI=10.1126/science.287.5459.1809;
RA   Tettelin H., Saunders N.J., Heidelberg J.F., Jeffries A.C.,
RA   Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F.,
RA   Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R.T., Peterson J.D.,
RA   Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,
RA   Dougherty B.A., Mason T.M., Ciecko A., Parksey D.S., Blair E.,
RA   Cittone H., Clark E.B., Cotton M.D., Utterback T.R., Khouri H.M.,
RA   Qin H., Vamathevan J.J., Gill J., Scarlato V., Masignani V., Pizza M.,
RA   Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R.,
RA   Venter J.C.;
RT   "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT   MC58.";
RL   Science 287:1809-1815(2000).
DR   EMBL; AE002098; AAF42439.1; -; Genomic_DNA.
DR   PIR; D81003; D81003.
DR   TIGR; NMB2131; -.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   71 AA;  8091 MW;  4A25DBB87621F66B CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 71;
  Best Local Similarity   40.0%;  Pred. No. 8.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FKLGVYTCLG 11


RESULT 46
Q8V1R5_9PAPI
ID   Q8V1R5_9PAPI PRELIMINARY;      PRT;    71 AA.
AC   Q8V1R5;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Major capsid protein (Fragment).
OS   Human papillomavirus.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   unclassified Papillomaviridae.
OX   NCBI_TaxID=10566;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22676373; PubMed=12791874;
RX   DOI=10.1128/JCM.41.6.2509-2514.2003;
RA   Antonsson A., Karanfilovska S., Lindqvist P.G., Hansson B.G.;
RT   "General acquisition of human papillomavirus infections of skin occurs
```

```
RT   in early infancy.";
RL   J. Clin. Microbiol. 41:2509-2514(2003).
DR   EMBL; AF455145; AAL57872.1; -; Genomic_DNA.
DR   HSSP; Q9WPH4; 1DZL.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR002210; PV_capsid_L1.
DR   Pfam; PF00500; Late_protein_L1; 1.
DR   PRINTS; PR00865; HPVCAPSIDL1.
DR   ProDom; PD000544; PV_capsid_L1; 1.
FT   NON_TER       1      1
FT   NON_TER      71     71
SQ   SEQUENCE   71 AA;  7998 MW;  8F7DDC13BB31393B CRC64;

  Query Match            100.0%;  Score 33;  DB 2;  Length 71;
  Best Local Similarity   40.0%;  Pred. No. 8.1e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         33 FIVGCAPCIG 42


RESULT 47
Q8IWZ9_HUMAN
ID   Q8IWZ9_HUMAN PRELIMINARY;      PRT;    72 AA.
AC   Q8IWZ9;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypertension associated protein (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Choi I., Moon K.D., Song I., Wiley J.W.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF519545; AAN77120.1; -; Genomic_DNA.
DR   SMR; Q8IWZ9; 1-71.
DR   Ensembl; ENSG00000111664; Homo sapiens.
DR   InterPro; IPR001680; WD40.
DR   Pfam; PF00400; WD40; 2.
DR   ProDom; PD000018; WD40; 1.
DR   SMART; SM00320; WD40; 1.
DR   PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
DR   PROSITE; PS50082; WD_REPEATS_2; 1.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   WD repeat.
FT   NON_TER       1      1
FT   NON_TER      72     72
SQ   SEQUENCE   72 AA;  7993 MW;  B61C9DEF13DBEC64 CRC64;

  Query Match            100.0%;  Score 33;  DB 2;  Length 72;
  Best Local Similarity   40.0%;  Pred. No. 8.2e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db          2 FPNGEAICTG 11


RESULT 48
Q8BC64_9PAPI
ID   Q8BC64_9PAPI PRELIMINARY;      PRT;    72 AA.
AC   Q8BC64;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Major capsid protein (Fragment).
OS   Human papillomavirus.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   unclassified Papillomaviridae.
OX   NCBI_TaxID=10566;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22695342; PubMed=12810883; DOI=10.1099/vir.0.18836-0;
RA   Antonsson A., Erfurt C., Hazard K., Holmgren V., Simon M., Kataoka A.,
RA   Hossain S., Hakangard C., Hansson B.G.;
RT   "Prevalence and type spectrum of human papillomaviruses in healthy
RT   skin samples collected in three continents.";
RL   J. Gen. Virol. 84:1881-1886(2003).
DR   EMBL; AF542100; AAN28675.1; -; Genomic_DNA.
DR   HSSP; Q9WPH4; 1DZL.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR002210; PV_capsid_L1.
DR   Pfam; PF00500; Late_protein_L1; 1.
DR   PRINTS; PR00865; HPVCAPSIDL1.
DR   ProDom; PD000544; PV_capsid_L1; 1.
FT   NON_TER       1      1
FT   NON_TER      72     72
SQ   SEQUENCE   72 AA;  8182 MW;  6765F1107AC17C82 CRC64;

  Query Match            100.0%;  Score 33;  DB 2;  Length 72;
  Best Local Similarity   40.0%;  Pred. No. 8.2e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         33 FIVGCTPCIG 42


RESULT 49
Q8JSL4_9PAPI
ID   Q8JSL4_9PAPI PRELIMINARY;      PRT;    72 AA.
AC   Q8JSL4;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Major capsid protein (Fragment).
OS   Chimpanzee papillomavirus.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   unclassified Papillomaviridae; primate papillomaviruses.
OX   NCBI_TaxID=203380;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22326806; PubMed=12438579;
RX   DOI=10.1128/JVI.76.24.12537-12542.2002;
RA   Antonsson A., Hansson B.G.;
RT   "Healthy skin of many animal species harbors papillomaviruses which
RT   are closely related to their human counterparts.";
RL   J. Virol. 76:12537-12542(2002).
DR   EMBL; AF488687; AAM97872.1; -; Genomic_DNA.
DR   HSSP; Q9WPH4; 1DZL.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR002210; PV_capsid_L1.
DR   Pfam; PF00500; Late_protein_L1; 1.
DR   PRINTS; PR00865; HPVCAPSIDL1.
DR   ProDom; PD000544; PV_capsid_L1; 1.
FT   NON_TER       1      1
FT   NON_TER      72     72
SQ   SEQUENCE   72 AA;  7953 MW;  CEBAA6474A7F27D0 CRC64;

  Query Match            100.0%;  Score 33;  DB 2;  Length 72;
  Best Local Similarity   40.0%;  Pred. No. 8.2e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         33 FIIGCTPCLG 42
```

```
RESULT 50
Q9Q2R1_9PAPI
ID   Q9Q2R1_9PAPI PRELIMINARY;      PRT;    72 AA.
AC   Q9Q2R1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Major capsid protein L1 (Fragment).
OS   Human papillomavirus.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   unclassified Papillomaviridae.
OX   NCBI_TaxID=10566;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=20541962; PubMed=11090162;
RX   DOI=10.1128/JVI.74.24.11636-11641.2000;
RA   Antonsson A., Forslund O., Ekberg H., Sterner G., Hansson B.G.;
RT   "The ubiquity and impressive genomic diversity of human skin
RT   papillomaviruses suggest a commensalic nature of these viruses.";
RL   J. Virol. 74:11636-11641(2000).
DR   EMBL; AF217681; AAF25471.1; -; Genomic_DNA.
DR   HSSP; Q9WPH4; 1DZL.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR002210; PV_capsid_L1.
DR   Pfam; PF00500; Late_protein_L1; 1.
DR   PRINTS; PR00865; HPVCAPSIDL1.
DR   ProDom; PD000544; PV_capsid_L1; 1.
FT   NON_TER       1      1
FT   NON_TER      72     72
SQ   SEQUENCE   72 AA;  8060 MW;  695D736F8F2CBA4B CRC64;

  Query Match             100.0%;  Score 33;  DB 2;  Length 72;
  Best Local Similarity   40.0%;  Pred. No. 8.2e+02;
  Matches    4;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FXXGXXXCXG 10
              |::|:::|:|
Db         33 FIVGCTPCIG 42


Search completed: March  8, 2006, 11:23:55
Job time : 236 secs
```

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.

- If you encounter an accession number from an older search run against UniProt (results file extension **.rup)** that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

  http://www.pir.uniprot.org/database/archive.shtml

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2006, 21:43:16 ; Search time 118.8 Seconds
(without alignments)
142.531 Million cell updates/sec

Title: US-10-751-235-10
Perfect score: 122
Sequence: 1 LVAEVSEFLFGSGFAIAEGPLWTA 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	122	100.0	539	2	Q6TBX7_ARATH	Q6tbx7 arabidopsis
2	122	100.0	552	2	Q8RWV4_ARATH	Q8rwv4 arabidopsis
3	118	96.7	566	2	Q9SCP8_ARATH	Q9scp8 arabidopsis
4	110	90.2	584	2	Q9AV27_ORYSA	Q9av27 oryza sativ
5	57	46.7	344	1	SYFA_BACSK	Q5wej5 bacillus cl
6	56	45.9	324	2	Q9CFC7_LACLA	Q9cfc7 lactococcus
7	56	45.9	344	1	SYFA_BACHD	Q9k895 bacillus ha
8	55.5	45.5	470	2	Q5FPI9_GLUOX	Q5fpi9 gluconobact
9	55	45.1	339	1	SYFA_GEOSL	Q74d00 geobacter s
10	55	45.1	350	2	Q8YJV8_ANASP	Q8yjv8 anabaena sp
11	54	44.3	593	2	Q9C6S0_ARATH	Q9c6s0 arabidopsis
12	54	44.3	595	2	Q93VK5_ARATH	Q93vk5 arabidopsis
13	52.5	43.0	97	2	Q8NQS2_CORGL	Q8nqs2 corynebacte
14	52	42.6	321	1	LPXK_RICCN	Q92gn1 rickettsia
15	52	42.6	321	1	LPXK_RICRI	P58187 rickettsia
16	52	42.6	321	2	Q7P9K9_RICSI	Q7p9k9 rickettsia
17	52	42.6	325	1	LPXK_RICMO	P58186 rickettsia
18	52	42.6	325	2	Q4UN11_RICFE	Q4un11 rickettsia
19	52	42.6	344	1	SYFA_GEOKA	Q5kwe4 geobacillus
20	51.5	42.2	68	2	Q6MJ94_BDEBA	Q6mj94 bdellovibri
21	51.5	42.2	296	2	Q4TSR2_9SPHN	Q4tsr2 erythrobact
22	51	41.8	339	1	SYFA_THETN	Q8r9c6 thermoanaer
23	51	41.8	375	2	Q7M860_WOLSU	Q7m860 wolinella s
24	51	41.8	1258	2	Q9HED9_NEUCR	Q9hed9 neurospora
25	51	41.8	1709	2	Q93H59_STRAW	Q93h59 streptomyce
26	50	41.0	297	2	Q6HCC7_BACHK	Q6hcc7 bacillus th
27	50	41.0	297	2	Q632T7_BACCZ	Q632t7 bacillus ce
28	50	41.0	350	1	SYFA_LISIN	Q92ci7 listeria in
29	50	41.0	350	1	SYFA_LISMF	Q720k8 listeria mo
30	50	41.0	350	1	SYFA_LISMO	Q8y7g2 listeria mo
31	50	41.0	601	2	Q4P0G2_USTMA	Q4p0g2 ustilago ma
32	50	41.0	670	2	Q5W270_9ENTR	Q5w270 serratia sp
33	50	41.0	718	2	Q7WAV9_BORPA	Q7wav9 bordetella
34	50	41.0	718	2	Q7WK16_BORBR	Q7wk16 bordetella
35	49.5	40.6	287	2	Q9KCQ8_BACHD	Q9kcq8 bacillus ha
36	49.5	40.6	524	2	Q89PW1_BRAJA	Q89pw1 bradyrhizob
37	49	40.2	216	2	Q5TWN3_ANOGA	Q5twn3 anopheles g
38	49	40.2	275	2	Q5SJ53_THET8	Q5sj53 thermus the
39	49	40.2	275	2	Q72JH5_THET2	Q72jh5 thermus the
40	49	40.2	287	2	Q9KD24_BACHD	Q9kd24 bacillus ha
41	49	40.2	444	2	Q28493_ARCFU	Q28493 archaeoglob
42	49	40.2	494	1	SYK_SULSO	P95970 sulfolobus
43	49	40.2	532	2	Q6C2L7_YARLI	Q6c2l7 yarrowia li
44	49	40.2	572	1	INDY1_DROME	Q9vvt2 drosophila
45	49	40.2	582	2	Q7QIT2_ANOGA	Q7qit2 anopheles g

ALIGNMENTS

RESULT 1
Q6TBX7_ARATH
ID Q6TBX7_ARATH PRELIMINARY; PRT; 539 AA.
AC Q6TBX7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Chloroplast carotenoid epsilon-ring hydroxylase.
GN Name=LUT1;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14709673; DOI=10.1073/pnas.2237237100;
RA Tian L., Musetti V., Kim J., Magallanes-Lundback M., DellaPenna D.;
RT "The Arabidopsis LUT1 locus encodes a member of the cytochrome p450
RT family that is required for carotenoid epsilon-ring hydroxylation
RT activity.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:402-407(2004).
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; AY424805; AAR83120.1; -; mRNA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002401; EP450I.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR00463; EP450I.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase;
KW Transmembrane.
SQ SEQUENCE 539 AA; 60555 MW; 4C25C728B676AEEB CRC64;

Query Match 100.0%; Score 122; DB 2; Length 539;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LVAEVSEFLFGSGFAIAEGPLWTA 24
||||||||||||||||||||||||
Db 144 LVAEVSEFLFGSGFAIAEGPLWTA 167

RESULT 2
Q8RWV4_ARATH
ID Q8RWV4_ARATH PRELIMINARY; PRT; 552 AA.
AC Q8RWV4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative cytochrome P450 (Fragment).

```
GN   Name=At3g53130;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC   rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA   Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA   Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA   Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA   Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA   Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA   Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA   Davis R.W., Ecker J.R., Theologis A.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR   EMBL; AY091083; AAM13903.1; -; mRNA.
DR   HSSP; P14779; 1JPZ.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   InterPro; IPR002401; EP450I.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00463; EP450I.
DR   PRINTS; PR00385; P450.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase;
KW   Transmembrane.
FT   NON_TER       1      1
SQ   SEQUENCE   552 AA;  62073 MW;  C23CF8498B5B8440 CRC64;

  Query Match             100.0%;  Score 122;  DB 2;  Length 552;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-09;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLWTA 24
              ||||||||||||||||||||||||
Db        157 LVAEVSEFLFGSGFAIAEGPLWTA 180


RESULT 3
Q9SCP8_ARATH
ID   Q9SCP8_ARATH PRELIMINARY;      PRT;   566 AA.
AC   Q9SCP8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Cytochrom P450-like protein.
GN   Name=T4D2.60;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC   rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Nyakatura G., Fartmann B., Dauner D., Sterr W., Holland R.,
RA   Weichselgartner M., Mewes H.W., Lemcke K., Mayer K.F.X., Quetier F.,
RA   Salanoubat M.;
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   EU Arabidopsis sequencing project;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR   EMBL; AL132958; CAB64216.1; -; Genomic_DNA.
DR   PIR; T46159; T46159.
DR   HSSP; P14779; 1JPZ.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   InterPro; IPR002401; EP450I.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00463; EP450I.
DR   PRINTS; PR00385; P450.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase;
KW   Transmembrane.
SQ   SEQUENCE   566 AA;  63596 MW;  B884E8996B1A4C7D CRC64;

  Query Match             96.7%;  Score 118;  DB 2;  Length 566;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-09;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLWT 23
              |||||||||||||||||||||||
Db        144 LVAEVSEFLFGSGFAIAEGPLWT 166


RESULT 4
Q9AV27_ORYSA
ID   Q9AV27_ORYSA PRELIMINARY;      PRT;   584 AA.
AC   Q9AV27; Q7XCI2;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Putative cytochrome P450 monooxygenase.
GN   ORFNames=OSJNBa0001O14.16;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Burr P.C.,
RA   Hsiao J., Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,
RA   Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,
RA   Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Buell R.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RA   The Rice Chromosome 10 Sequencing Consortium;
RT   "In-depth view of structure, activity, and evolution of rice
RT   chromosome 10.";
RL   Science 300:1566-1569(2003).
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RA   Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR   EMBL; AE017117; AAP54891.1; -; Genomic_DNA.
DR   EMBL; AC025783; AAK20054.1; -; Genomic_DNA.
DR   HSSP; P14779; 1JPZ.
DR   Gramene; Q9AV27; -.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001128; Cytochrome_P450.
DR   InterPro; IPR002401; EP450I.
DR   Pfam; PF00067; p450; 1.
DR   PRINTS; PR00463; EP450I.
DR   PRINTS; PR00385; P450.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase;
KW   Transmembrane.
SQ   SEQUENCE   584 AA;  64798 MW;  1A55160A532DCE83 CRC64;
```

```
  Query Match           90.2%;  Score 110;  DB 2;  Length 584;
  Best Local Similarity  95.7%;  Pred. No. 8e-08;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         1 LVAEVSEFLFGSGFAIAEGPLWT 23
             |||||||||||||||||||| |||
Db       152 LVAEVSEFLFGSGFAIAEGALWT 174


RESULT 5
SYFA_BACSK
ID   SYFA_BACSK     STANDARD;      PRT;   344 AA.
AC   Q5WEJ5;
DT   10-MAY-2005 (Rel. 47, Created)
DT   10-MAY-2005 (Rel. 47, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20)
DE   (Phenylalanine--tRNA ligase alpha chain) (PheRS).
GN   Name=pheS; OrderedLocusNames=ABC2680;
OS   Bacillus clausii (strain KSM-K16).
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=66692;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA   Takaki Y., Kageyama Y., Shimamura S., Suzuki H., Nishi S., Hatada Y.,
RA   Kawai S., Ito S., Horikoshi K.;
RT   "The complete genome sequence of the alkaliphilic Bacillus clausii
RT   KSM-K16.";
RL   Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +
CC       diphosphate + L-phenylalanyl-tRNA(Phe).
CC   -!- COFACTOR: Binds 2 magnesium ions per tetramer (By similarity).
CC   -!- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Belongs to the class-II aminoacyl-tRNA synthetase
CC       family. Phe-tRNA synthetase alpha chain type 1 subfamily.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AP006627; BAD65215.1; -; Genomic_DNA.
DR   HAMAP; MF_00281; -; 1.
DR   InterPro; IPR004188; Phe_tRNA_synt_N.
DR   InterPro; IPR004529; PheS.
DR   InterPro; IPR002319; tRNA-synt_2d.
DR   InterPro; IPR006195; tRNA_ligase_II.
DR   PANTHER; PTHR11538; tRNA-synt_2d; 1.
DR   Pfam; PF02912; Phe_tRNA-synt_N; 1.
DR   Pfam; PF01409; tRNA-synt_2d; 1.
DR   TIGRFAMs; TIGR00468; pheS; 1.
DR   PROSITE; PS50862; AA_TRNA_LIGASE_II; 1.
KW   Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
KW   Magnesium; Metal-binding; Nucleotide-binding; Protein biosynthesis.
FT   METAL       256    256       Magnesium (By similarity).
SQ   SEQUENCE   344 AA;  38634 MW;  15F66E8B3556E1CC CRC64;

  Query Match           46.7%;  Score 57;  DB 1;  Length 344;
  Best Local Similarity  50.0%;  Pred. No. 5;
  Matches   10;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         1 LVAEVSEFLFGSGFAIAEGP 20
             :| |: |   | ||:| |||
Db       113 VVTEIEEIFLGLGFSIGEGP 132


RESULT 6
Q9CFC7_LACLA
ID   Q9CFC7_LACLA PRELIMINARY;      PRT;   324 AA.
AC   Q9CFC7;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Oxidoreductase.
GN   Name=ypjF; OrderedLocusNames=LL1554;
OS   Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX   NCBI_TaxID=1360;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=IL1403;
RX   MEDLINE=21235186; PubMed=11337471; DOI=10.1101/gr.GR-1697R;
RA   Bolotin A., Wincker P., Mauger S., Jaillon O., Malarme K.,
RA   Weissenbach J., Ehrlich S.D., Sorokin A.;
RT   "The complete genome sequence of the lactic acid bacterium Lactococcus
RT   lactis ssp. lactis IL1403.";
RL   Genome Res. 11:731-753(2001).
DR   EMBL; AE006386; AAK05652.1; -; Genomic_DNA.
DR   PIR; B86819; B86819.
DR   GO; GO:0016491; F:oxidoreductase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR000683; GFO/IDH/MocA_N.
DR   InterPro; IPR004104; GFO_IDH_MocA_C.
DR   Pfam; PF01408; GFO_IDH_MocA; 1.
DR   Pfam; PF02894; GFO_IDH_MocA_C; 1.
KW   Complete proteome.
SQ   SEQUENCE   324 AA;  36603 MW;  DBF47D6684EB259C CRC64;

  Query Match           45.9%;  Score 56;  DB 2;  Length 324;
  Best Local Similarity  50.0%;  Pred. No. 6.8;
  Matches   10;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         5 VSEFLFGSGFAIAEGPLWTA 24
             | :||||:||  |:|  |::
Db       182 VLQFLFGTGFDAAKGIYWSS 201


RESULT 7
SYFA_BACHD
ID   SYFA_BACHD     STANDARD;      PRT;   344 AA.
AC   Q9K895;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20)
DE   (Phenylalanine--tRNA ligase alpha chain) (PheRS).
GN   Name=pheS; OrderedLocusNames=BH3111;
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
CC   -!- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +
CC       diphosphate + L-phenylalanyl-tRNA(Phe).
CC   -!- COFACTOR: Binds 2 magnesium ions per tetramer (By similarity).
CC   -!- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Belongs to the class-II aminoacyl-tRNA synthetase
CC       family. Phe-tRNA synthetase alpha chain type 1 subfamily.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------------
DR   EMBL; BA000004; BAB06830.1; -; Genomic_DNA.
DR   PIR; G84038; G84038.
DR   HSSP; P27001; 1JJC.
DR   HAMAP; MF_00281; -; 1.
DR   InterPro; IPR004188; Phe_tRNA_synt_N.
DR   InterPro; IPR004529; PheS.
DR   InterPro; IPR002319; tRNA-synt_2d.
DR   InterPro; IPR006195; tRNA_ligase_II.
DR   PANTHER; PTHR11538; tRNA-synt_2d; 1.
DR   Pfam; PF02912; Phe_tRNA-synt_N; 1.
DR   Pfam; PF01409; tRNA-synt_2d; 1.
DR   TIGRFAMs; TIGR00468; pheS; 1.
DR   PROSITE; PS50862; AA_TRNA_LIGASE_II; 1.
KW   Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
KW   Magnesium; Metal-binding; Nucleotide-binding; Protein biosynthesis.
FT   METAL       256    256       Magnesium (By similarity).
SQ   SEQUENCE   344 AA;  38931 MW;  498A039538E7A94C CRC64;

  Query Match             45.9%;  Score 56;  DB 1;  Length 344;
  Best Local Similarity   47.4%;  Pred. No. 7.2;
  Matches    9;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 VAEVSEFLFGSGFAIAEGP 20
              : || :   | ||:::|||
Db        114 ITEVEDLFIGLGFSVAEGP 132


RESULT 8
Q5FPI9_GLUOX
ID   Q5FPI9_GLUOX PRELIMINARY;      PRT;   470 AA.
AC   Q5FPI9;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Galactose-proton symporter.
GN   OrderedLocusNames=GOX1971;
OS   Gluconobacter oxydans (Gluconobacter suboxydans).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC   Acetobacteraceae; Gluconobacter.
OX   NCBI_TaxID=442;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=621H;
RX   PubMed=15665824; DOI=10.1038/nbt1062;
RA   Prust C., Hoffmeister M., Liesegang H., Wiezer A., Fricke W.F.,
RA   Ehrenreich A., Gottschalk G., Deppenmeier U.;
RT   "Complete genome sequence of the acetic acid bacterium Gluconobacter
RT   oxydans.";
RL   Nat. Biotechnol. 23:195-200(2005).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR   EMBL; CP000009; AAW61707.1; -; Genomic_DNA.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005351; F:sugar porter activity; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0008643; P:carbohydrate transport; IEA.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR005829; Sug_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   Pfam; PF00083; Sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW   Complete proteome; Transmembrane; Transport.
SQ   SEQUENCE   470 AA;  49835 MW;  ECBAC911C986E16C CRC64;

  Query Match             45.5%;  Score 55.5;  DB 2;  Length 470;
  Best Local Similarity   75.0%;  Pred. No. 11;
  Matches   12;  Conservative    1;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          9 LFGSGFAIAEGPL-WT 23
              || :|||| |||| ||
Db        365 LFVAGFAIGEGPLVWT 380


RESULT 9
SYFA_GEOSL
ID   SYFA_GEOSL     STANDARD;      PRT;   339 AA.
AC   Q74D00;
DT   10-MAY-2005 (Rel. 47, Created)
DT   10-MAY-2005 (Rel. 47, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20)
DE   (Phenylalanine--tRNA ligase alpha chain) (PheRS).
GN   Name=pheS; OrderedLocusNames=GSU1519;
OS   Geobacter sulfurreducens.
OC   Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
OC   Geobacteraceae; Geobacter.
OX   NCBI_TaxID=35554;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=PCA / ATCC 51573;
RX   PubMed=14671304; DOI=10.1126/science.1088727;
RA   Methe B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,
RA   Heidelberg J.F., Wu D., Wu M., Ward N.L., Beanan M.J., Dodson R.J.,
RA   Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,
RA   Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,
RA   Davidsen T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,
RA   Weidman J.F., Khouri H.M., Feldblyum T.V., Utterback T.R.,
RA   Van Aken S.E., Lovley D.R., Fraser C.M.;
RT   "Genome of Geobacter sulfurreducens: metal reduction in subsurface
RT   environments.";
RL   Science 302:1967-1969(2003).
CC   -!- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +
CC       diphosphate + L-phenylalanyl-tRNA(Phe).
CC   -!- COFACTOR: Binds 2 magnesium ions per tetramer (By similarity).
CC   -!- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Belongs to the class-II aminoacyl-tRNA synthetase
CC       family. Phe-tRNA synthetase alpha chain type 1 subfamily.
CC   ---------------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------------
DR   EMBL; AE017180; AAR34893.1; -; Genomic_DNA.
DR   HSSP; P27001; 1B7Y.
DR   TIGR; GSU1519; -.
DR   HAMAP; MF_00281; -; 1.
DR   InterPro; IPR004188; Phe_tRNA_synt_N.
DR   InterPro; IPR004529; PheS.
DR   InterPro; IPR002319; tRNA-synt_2d.
DR   InterPro; IPR006195; tRNA_ligase_II.
DR   PANTHER; PTHR11538; tRNA-synt_2d; 1.
DR   Pfam; PF02912; Phe_tRNA-synt_N; 1.
DR   Pfam; PF01409; tRNA-synt_2d; 1.
DR   TIGRFAMs; TIGR00468; pheS; 1.
DR   PROSITE; PS50862; AA_TRNA_LIGASE_II; 1.
KW   Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
KW   Magnesium; Metal-binding; Nucleotide-binding; Protein biosynthesis.
FT   METAL       253    253       Magnesium (By similarity).
SQ   SEQUENCE   339 AA;  37888 MW;  3A36C6EDC3BC5F6A CRC64;

  Query Match             45.1%;  Score 55;  DB 1;  Length 339;
  Best Local Similarity   45.0%;  Pred. No. 10;
  Matches    9;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;
```

```
Qy          1 LVAEVSEFLFGSGFAIAEGP 20
              :: |:|:   | || :||||
Db        113 VIEEISDIFAGLGFQVAEGP 132
```

RESULT 10
Q8YJV8_ANASP
ID Q8YJV8_ANASP PRELIMINARY; PRT; 350 AA.
AC Q8YJV8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Alr9005 protein.
GN OrderedLocusNames=alr9005;
OS Anabaena sp. (strain PCC 7120).
OG Plasmid pCC7120epsilon.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003605; BAB77491.1; -; Genomic_DNA.
DR PIR; AF2572; AF2572.
KW Complete proteome.
SQ SEQUENCE 350 AA; 39119 MW; 0AB1C62658675201 CRC64;

  Query Match 45.1%; Score 55; DB 2; Length 350;
  Best Local Similarity 61.1%; Pred. No. 10;
  Matches 11; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy          5 VSEFLFGSGFAIAEGPLW 22
              | | |   || ||| |||
Db        313 VDEALKAKGFVIAENPLW 330
```

RESULT 11
Q9C6S0_ARATH
ID Q9C6S0_ARATH PRELIMINARY; PRT; 593 AA.
AC Q9C6S0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome P450, putative.
GN Name=F5M6.19;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; AC079041; AAG50718.1; -; Genomic_DNA.
DR PIR; F86441; F86441.
DR HSSP; P14779; 1P0W.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002401; EP450I.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR00463; EP450I.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase;
KW Transmembrane.
SQ SEQUENCE 593 AA; 66643 MW; E80CBE9B8B2BD199 CRC64;

  Query Match 44.3%; Score 54; DB 2; Length 593;
  Best Local Similarity 36.4%; Pred. No. 24;
  Matches 8; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

```
Qy          1 LVAEVSEFLFGSGFAIAEGPLW 22
              ::||: :|: | |   |:| :|
Db        174 ILAEILDFVMGKGLIPADGEIW 195
```

RESULT 12
Q93VK5_ARATH
ID Q93VK5_ARATH PRELIMINARY; PRT; 595 AA.
AC Q93VK5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE At1g31800/68069_m00159.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Kim C.J., Chen H., Cheuk R., Shinn P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
RA Hsuan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
RA Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; AY058173; AAL25587.1; -; mRNA.
DR EMBL; AY056446; AAL08302.1; -; mRNA.
DR EMBL; AY142017; AAM98281.1; -; mRNA.
DR HSSP; P14779; 1P0W.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002401; EP450I.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR00463; EP450I.

```
DR   PRINTS; PR00385; P450.
DR   PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW   Heme; Iron; Metal-binding; Monooxygenase; Oxidoreductase;
KW   Transmembrane.
SQ   SEQUENCE   595 AA;  66846 MW;  A548A33A48323B7D CRC64;

  Query Match             44.3%;  Score 54;  DB 2;  Length 595;
  Best Local Similarity   36.4%;  Pred. No. 24;
  Matches    8;  Conservative    7;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLW 22
              ::||: :|: | |   |:| :|
Db        176 ILAEILDFVMGKGLIPADGEIW 197


RESULT 13
Q8NQS2_CORGL
ID   Q8NQS2_CORGL PRELIMINARY;      PRT;    97 AA.
AC   Q8NQS2; Q6M5K5;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical membrane protein (Putative membrane protein).
GN   OrderedLocusNames=Cgl1352, cg1524;
OS   Corynebacterium glutamicum (Brevibacterium flavum).
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX   NCBI_TaxID=1718;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA   Nakagawa S.;
RT   "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX   PubMed=12948626; DOI=10.1016/S0168-1656(03)00154-8;
RA   Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,
RA   Burkovski A., Dusch N., Eggeling L., Eikmanns B.J., Gaigalat L.,
RA   Goesmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
RA   McHardy A.C., Meyer F., Moeckel B., Pfefferle W., Puehler A.,
RA   Rey D.A., Rueckert C., Rupp O., Sahm H., Wendisch V.F., Wiegraebe I.,
RA   Tauch A.;
RT   "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
RT   and its impact on the production of L-aspartate-derived amino acids
RT   and vitamins.";
RL   J. Biotechnol. 104:5-25(2003).
DR   EMBL; BA000036; BAB98745.1; -; Genomic_DNA.
DR   EMBL; BX927152; CAF21362.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   97 AA;  10570 MW;  14FAB3646A259D46 CRC64;

  Query Match             43.0%;  Score 52.5;  DB 2;  Length 97;
  Best Local Similarity   57.7%;  Pred. No. 7.3;
  Matches   15;  Conservative    1;  Mismatches    7;  Indels    3;  Gaps    1;

Qy          1 LVAEVSEFL---FGSGFAIAEGPLWT 23
              |:| ||  |   |||| |||   |||
Db          9 LLAMVSTALRIRFGSGVAIAATVLWT 34


RESULT 14
LPXK_RICCN
ID   LPXK_RICCN     STANDARD;      PRT;   321 AA.
AC   Q92GN1;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Tetraacyldisaccharide 4'-kinase (EC 2.7.1.130) (Lipid A 4'-kinase).
GN   Name=lpxK; OrderedLocusNames=RC1092;
OS   Rickettsia conorii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia; spotted fever group.
OX   NCBI_TaxID=781;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=Malish 7;
RX   MEDLINE=21442074; PubMed=11557893; DOI=10.1126/science.1061471;
RA   Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA   Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA   Raoult D.;
RT   "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL   Science 293:2093-2098(2001).
CC   -!- FUNCTION: Transfers the gamma-phosphate of ATP to the 4'-position
CC       of a tetraacyldisaccharide 1-phosphate intermediate (termed DS-1-
CC       P) to form tetraacyldisaccharide 1,4'-bis-phosphate (lipid IVA)
CC       (By similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + (2-N,3-O-bis(3-hydroxytetradecanoyl)-
CC       beta-D-glucosaminyl)-(1->6)-(2-N,3-O-bis(3-hydroxytetradecanoyl)-
CC       beta-D-glucosaminyl phosphate) = ADP + (2-N,3-O-bis(3-
CC       hydroxytetradecanoyl)-4-O-phosphono-beta-D-glucosaminyl)-(1->6)-
CC       (2-N,3-O-bis(3-hydroxytetradecanoyl)-beta-D-glucosaminyl
CC       phosphate).
CC   -!- PATHWAY: Lipid A biosynthesis; sixth step.
CC   -!- SIMILARITY: Belongs to the lpxK family.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; AE008660; AAL03630.1; -; Genomic_DNA.
DR   PIR; D97836; D97836.
DR   HAMAP; MF_00409; -; 1.
DR   InterPro; IPR003758; LpxK.
DR   Pfam; PF02606; LpxK; 1.
DR   TIGRFAMs; TIGR00682; lpxK; 1.
KW   ATP-binding; Complete proteome; Kinase; Lipid A biosynthesis;
KW   Lipid synthesis; Nucleotide-binding; Transferase.
FT   NP_BIND      54     61       ATP (Potential).
SQ   SEQUENCE   321 AA;  36059 MW;  9FF4672CE1CF99E2 CRC64;

  Query Match             42.6%;  Score 52;  DB 1;  Length 321;
  Best Local Similarity   52.4%;  Pred. No. 27;
  Matches   11;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPL 21
              :|:  |: |||:|| |  |||
Db        155 IVSVDSQRLFGNGFLIPAGPL 175


RESULT 15
LPXK_RICRI
ID   LPXK_RICRI     STANDARD;      PRT;   321 AA.
AC   P58187;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Tetraacyldisaccharide 4'-kinase (EC 2.7.1.130) (Lipid A 4'-kinase).
GN   Name=lpxK;
OS   Rickettsia rickettsii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia; spotted fever group.
OX   NCBI_TaxID=783;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC   STRAIN=84-21C;
RX   MEDLINE=21219194; PubMed=11319266;
RA   Andersson J.O., Andersson S.G.E.;
RT   "Pseudogenes, junk DNA, and the dynamics of Rickettsia genomes.";
RL   Mol. Biol. Evol. 18:829-839(2001).
CC   -!- FUNCTION: Transfers the gamma-phosphate of ATP to the 4'-position
CC       of a tetraacyldisaccharide 1-phosphate intermediate (termed DS-1-
CC       P) to form tetraacyldisaccharide 1,4'-bis-phosphate (lipid IVA)
```

```
CC       (By similarity).
CC    -!- CATALYTIC ACTIVITY: ATP + (2-N,3-O-bis(3-hydroxytetradecanoyl)-
CC        beta-D-glucosaminyl)-(1->6)-(2-N,3-O-bis(3-hydroxytetradecanoyl)-
CC        beta-D-glucosaminyl phosphate) = ADP + (2-N,3-O-bis(3-
CC        hydroxytetradecanoyl)-4-O-phosphono-beta-D-glucosaminyl)-(1->6)-
CC        (2-N,3-O-bis(3-hydroxytetradecanoyl)-beta-D-glucosaminyl
CC        phosphate).
CC    -!- PATHWAY: Lipid A biosynthesis; sixth step.
CC    -!- SIMILARITY: Belongs to the lpxK family.
CC    ----------------------------------------------------------------------
CC    This Swiss-Prot entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use as long as its content is in no way modified and this statement is not
CC    removed.
CC    ----------------------------------------------------------------------
DR    EMBL; AJ293329; CAC33716.1; -; Genomic_DNA.
DR    HAMAP; MF_00409; -; 1.
DR    InterPro; IPR003758; LpxK.
DR    Pfam; PF02606; LpxK; 1.
DR    TIGRFAMs; TIGR00682; lpxK; 1.
KW    ATP-binding; Kinase; Lipid A biosynthesis; Lipid synthesis;
KW    Nucleotide-binding; Transferase.
FT    NP_BIND      54     61       ATP (Potential).
SQ    SEQUENCE   321 AA;  36083 MW;  45E6CEB6CC50C0AE CRC64;

  Query Match             42.6%;  Score 52;  DB 1;  Length 321;
  Best Local Similarity   52.4%;  Pred. No. 27;
  Matches   11;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPL 21
              :|:  |: |||:|| |  |||
Db        155 IVSVDSQRLFGNGFLIPAGPL 175


Search completed: March  7, 2006, 21:58:09
Job time : 122.8 secs
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2006, 21:43:03 ; Search time 234.8 Seconds
(without alignments)
44.911 Million cell updates/sec

Title: US-10-751-235-10
Perfect score: 122
Sequence: 1 LVAEVSEFLFGSGFAIAEGPLWTA 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	122	100.0	24	9	AEB16912	Aeb16912 Cytochrom
2	122	100.0	539	9	AEB16908	Aeb16908 Thale cre
3	122	100.0	560	8	ADT56240	Adt56240 Plant pol
4	114	93.4	208	9	AEB16923	Aeb16923 Sunflower
5	110	90.2	362	9	AEB16920	Aeb16920 Wheat cyt
6	110	90.2	382	8	ADY22850	Ady22850 Plant ful
7	110	90.2	545	9	AEB16919	Aeb16919 Barley cy
8	110	90.2	561	9	AEB16918	Aeb16918 Rice cyto
9	110	90.2	588	9	AEB16954	Aeb16954 Rice cyto
10	59	48.4	167	6	ABU25025	Abu25025 Protein e
11	59	48.4	342	6	ABU25075	Abu25075 Protein e
12	58	47.5	632	9	AEB16936	Aeb16936 Rice cyto
13	58	47.5	662	8	ADX67984	Adx67984 Plant ful
14	57	46.7	342	9	AEB16939	Aeb16939 Wheat cyt
15	57	46.7	508	9	AEB16937	Aeb16937 Barley cy
16	56	45.9	324	5	ABB54899	Abb54899 Lactococc
17	56	45.9	344	8	ADS28360	Ads28360 Bacterial
18	54	44.3	579	9	AEB16940	Aeb16940 Tomato cy
19	54	44.3	595	9	AEB16935	Aeb16935 Thale cre
20	52.5	43.0	97	4	AAG91239	Aag91239 C glutami
21	50	41.0	350	5	ABB49267	Abb49267 Listeria
22	50	41.0	350	6	ABU33039	Abu33039 Protein e
23	49.5	40.6	287	8	ADS28192	Ads28192 Bacterial
24	49	40.2	360	8	ADN25200	Adn25200 Bacterial
25	49	40.2	572	4	ABB60315	Abb60315 Drosophil
26	49	40.2	572	4	ABB66903	Abb66903 Drosophil
27	49	40.2	572	5	ABB79611	Abb79611 Drosophil
28	49	40.2	572	6	ABR40099	Abr40099 dIndyaa s
29	49	40.2	572	8	ADP64789	Adp64789 Drosophil
30	48	39.3	173	6	ABU20163	Abu20163 Protein e
31	48	39.3	303	8	ADS21348	Ads21348 Bacterial
32	48	39.3	316	9	ABM91222	Abm91222 M. xanthu
33	48	39.3	394	8	ADX76358	Adx76358 Plant ful
34	48	39.3	483	8	ADT87787	Adt87787 Plant hom
35	48	39.3	523	7	ADC61143	Adc61143 Baeyer-Vi
36	48	39.3	1032	8	ADJ49445	Adj49445 Oil-assoc
37	48	39.3	1124	8	ADN20724	Adn20724 Bacterial
38	47.5	38.9	103	3	AAB43990	Aab43990 Human can
39	47	38.5	168	4	AAU29352	Aau29352 Novel mar
40	47	38.5	255	8	ADX88517	Adx88517 Plant ful
41	47	38.5	326	8	ADS42528	Ads42528 Bacterial
42	47	38.5	359	8	ADS22754	Ads22754 Bacterial
43	47	38.5	489	8	ADX80107	Adx80107 Plant ful
44	47	38.5	512	8	ABM84611	Abm84611 Human dia
45	47	38.5	683	7	ABO79928	Abo79928 Pseudomon

ALIGNMENTS

RESULT 1
AEB16912
ID AEB16912 standard; peptide; 24 AA.
XX
AC AEB16912;
XX
DT 08-SEP-2005 (first entry)
XX
DE Cytochrome P450 monooxygenase transmembrane domain peptide SEQ ID NO: 10.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant;
KW cytochrome P450.
XX
OS Unidentified.
XX
PN US2005150002-A1.
XX
PD 07-JUL-2005.
XX
PF 02-JAN-2004; 2004US-00751235.
XX
PR 02-JAN-2004; 2004US-00751235.
XX
PA (DELL/) DELLAPENNA D.
PA (TIAN/) TIAN L.
PA (KIMJ/) KIM J.
XX
PI Dellapenna D, Tian L, Kim J;
XX
DR WPI; 2005-487984/49.
XX
PT New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.
XX
PS Claim 6; SEQ ID NO 10; 135pp; English.
XX
CC The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is a cytochrome P450 monooxygenase
CC conserved transmembrane domain peptide.
XX

```
SQ   Sequence 24 AA;

  Query Match             100.0%;  Score 122;  DB 9;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-12;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLWTA 24
              ||||||||||||||||||||||||
Db          1 LVAEVSEFLFGSGFAIAEGPLWTA 24


RESULT 2
AEB16908
ID   AEB16908 standard; protein; 539 AA.
XX
AC   AEB16908;
XX
DT   08-SEP-2005  (first entry)
XX
DE   Thale cress LUT1 (cytochrome P450 97C1) protein, SEQ ID NO: 4.
XX
KW   Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C1;
KW   cytochrome P450 97C1.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .36
FT                   /note= "Chloroplast-targeting N-terminal transit peptide"
FT   Misc-difference 166. .167
FT                   /note= "Encoded by bases of positions 496 to 650"
FT   Misc-difference 265. .266
FT                   /note= "Encoded by bases of positions 942 to 1037"
FT   Misc-difference 332. .333
FT                   /note= "Encoded by bases of positions 1242 to 1326"
FT   Region          347. .352
FT                   /note= "Molecular oxygen binding pocket"
FT   Misc-difference 365. .366
FT                   /note= "Encoded by bases of positions 1420 to 1523"
FT   Misc-difference 413. .414
FT                   /note= "Encoded by bases of positions 1662 to 1841"
FT   Misc-difference 449. .450
FT                   /note= "Encoded by bases of positions 1839 to 12005"
FT   Misc-difference 476. .477
FT                   /note= "Encoded by bases of positions 2081 to 2183"
FT   Region          480. .489
FT                   /note= "Cysteine motif region"
FT   Misc-difference 530. .531
FT                   /note= "Encoded by bases of positions 2340 to 2440"
XX
PN   US2005150002-A1.
XX
PD   07-JUL-2005.
XX
PF   02-JAN-2004; 2004US-00751235.
XX
PR   02-JAN-2004; 2004US-00751235.
XX
PA   (DELL/) DELLAPENNA D.
PA   (TIAN/) TIAN L.
PA   (KIMJ/) KIM J.
XX
PI   Dellapenna D,  Tian L,  Kim J;
XX
DR   WPI; 2005-487984/49.
XX
PT   New expression vector comprising a nucleic acid sequence encoding a
PT   polypeptide having monooxygenase P450 activity, useful in altering the
PT   carotenoid production in a plant for enhancing production of specific
PT   carotenoid compounds.
XX
PS   Claim 9; SEQ ID NO 4; 135pp; English.
XX
CC   The present invention relates to genes, proteins and methods comprising
CC   carotenoid monooxygenases in the cytochrome P450 family. The invention
CC   also relates to altering carotenoid ratios in plants and microorganisms
CC   using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC   invention is useful in altering the carotenoid production in a plant for
CC   enhancing production of specific carotenoid compounds that are potent
CC   antioxidants. The present sequence is thale cress LUT1 [cytochrome P450
CC   monooxygenase (CYP97C1); At3g53130 gene] protein. Note: The current
CC   sequence is that of thale cress LUT1 protein that is encoded by cDNA
CC   located on chromosome 3.
XX
SQ   Sequence 539 AA;

  Query Match             100.0%;  Score 122;  DB 9;  Length 539;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-11;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLWTA 24
              ||||||||||||||||||||||||
Db        144 LVAEVSEFLFGSGFAIAEGPLWTA 167


RESULT 3
ADT56240
ID   ADT56240 standard; protein; 560 AA.
XX
AC   ADT56240;
XX
DT   13-JAN-2005  (first entry)
XX
DE   Plant polypeptide, SEQ ID 6317.
XX
KW   Plant; transgenic; cold tolerance; growth rate; drought tolerance;
KW   disease resistance; galactomannan production; plant growth regulator;
KW   heat tolerance; herbicide tolerance; lignin production;
KW   extreme osmotic condition tolerance; pathogens resistance;
KW   pest resistance; yield improvement; seed oil yield; seed protein yield.
XX
OS   Viridiplantae.
XX
PN   US2004216190-A1.
XX
PD   28-OCT-2004.
XX
PF   18-DEC-2003; 2003US-00739930.
XX
PR   28-APR-2003; 2003US-00424599.
PR   28-APR-2003; 2003US-00425115.
XX
PA   (KOVA/) KOVALIC D K.
XX
PI   Kovalic DK;
XX
DR   WPI; 2004-757369/74.
XX
PT   New recombinant DNA constructs useful in the field of biochemistry and
PT   genetics, and in particular for producing transgenic plants with improved
PT   biological characteristics.
XX
PS   Claim 2; SEQ ID NO 6317; 14pp; English.
XX
CC   The invention relates a recombinant DNA construct comprising a
CC   polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
CC   1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC   (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
CC   Arabidopsis, wheat and rape but the specification does not indicate which
CC   sequences is derived from which organism. Also included is a method of
CC   producing a plant having an improved property, comprising transforming a
CC   plant with a recombinant DNA construct comprising a promoter region
CC   functional in a plant cell operably joined to a polynucleotide encoding a
CC   polypeptide associated with the property, and growing the transformed
CC   plant. The property is selected from improving plant cold tolerance, for
CC   manipulating growth rate in plant cells by modification of the cell cycle
```

CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant disease, for galactomannan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and
CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to extreme osmotic and drought conditions, and improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant protein
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040216190.
XX
SQ Sequence 560 AA;

Query Match 100.0%; Score 122; DB 8; Length 560;
Best Local Similarity 100.0%; Pred. No. 5.6e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LVAEVSEFLFGSGFAIAEGPLWTA 24
||||||||||||||||||||||||
Db 165 LVAEVSEFLFGSGFAIAEGPLWTA 188

RESULT 4
AEB16923
ID AEB16923 standard; protein; 208 AA.
XX
AC AEB16923;
XX
DT 08-SEP-2005 (first entry)
XX
DE Sunflower cytochrome P450 97C (CYP97C) protein, SEQ ID NO: 21.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C;
KW cytochrome P450 97C.
XX
OS Helianthus annuus.
XX
PN US2005150002-A1.
XX
PD 07-JUL-2005.
XX
PF 02-JAN-2004; 2004US-00751235.
XX
PR 02-JAN-2004; 2004US-00751235.
XX
PA (DELL/) DELLAPENNA D.
PA (TIAN/) TIAN L.
PA (KIMJ/) KIM J.
XX
PI Dellapenna D, Tian L, Kim J;
XX
DR WPI; 2005-487984/49.
DR N-PSDB; AEB16929.
DR GENBANK; BQ971938.
XX
PT New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.
XX
PS Claim 9; SEQ ID NO 21; 135pp; English.
XX
CC The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is sunflower cytochrome P450
CC monooxygenase (CYP97C) protein.
XX
SQ Sequence 208 AA;

Query Match 93.4%; Score 114; DB 9; Length 208;
Best Local Similarity 95.8%; Pred. No. 3.6e-10;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LVAEVSEFLFGSGFAIAEGPLWTA 24
||||||||||||||||||||| ||||
Db 29 LVAEVSEFLFGSGFAIAEGSLWTA 52

RESULT 5
AEB16920
ID AEB16920 standard; protein; 362 AA.
XX
AC AEB16920;
XX
DT 08-SEP-2005 (first entry)
XX
DE Wheat cytochrome P450 97C (CYP97C) protein, SEQ ID NO: 18.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C;
KW cytochrome P450 97C.
XX
OS Triticum aestivum.
XX
PN US2005150002-A1.
XX
PD 07-JUL-2005.
XX
PF 02-JAN-2004; 2004US-00751235.
XX
PR 02-JAN-2004; 2004US-00751235.
XX
PA (DELL/) DELLAPENNA D.
PA (TIAN/) TIAN L.
PA (KIMJ/) KIM J.
XX
PI Dellapenna D, Tian L, Kim J;
XX
DR WPI; 2005-487984/49.
DR N-PSDB; AEB16926.
DR GENBANK; CA497665, BG906289, CA742365, CA742792.
XX
PT New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.
XX
PS Claim 9; SEQ ID NO 18; 135pp; English.
XX
CC The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is wheat cytochrome P450 monooxygenase
CC (CYP97C) protein.
XX

```
SQ   Sequence 362 AA;

  Query Match             90.2%;  Score 110;  DB 9;  Length 362;
  Best Local Similarity   95.7%;  Pred. No. 3.2e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLWT 23
              ||||||||||||||||||| |||
Db        141 LVAEVSEFLFGSGFAIAEGALWT 163


RESULT 6
ADY22850
ID   ADY22850 standard; protein; 382 AA.
XX
AC   ADY22850;
XX
DT   21-APR-2005  (first entry)
XX
DE   Plant full length insert polypeptide seqid 70634.
XX
KW   plant protectant; plant growth regulant; gene therapy; plant;
KW   recombinant DNA construct; physical array; plant breeding marker;
KW   cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW   extreme osmotic condition; pathogen tolerance; pest tolerance;
KW   growth rate; cell cycle pathway; disease resistance;
KW   galactomannan production; lignin production; plant growth regulator;
KW   yield; plant growth; plant development; seed oil; protein yield;
KW   protein content.
XX
OS   Unidentified.
XX
PN   US2004034888-A1.
XX
PD   19-FEB-2004.
XX
PF   28-APR-2003; 2003US-00425114.
XX
PR   06-MAY-1999;   99US-00304517.
PR   05-NOV-2001; 2001US-00985678.
XX
PA   (LIUJ/) LIU J.
PA   (ZHOU/) ZHOU Y.
PA   (KOVA/) KOVALIC D K.
PA   (SCRE/) SCREEN S E.
PA   (TABA/) TABASKA J E.
PA   (CAOY/) CAO Y.
XX
PI   Liu J,  Zhou Y,  Kovalic DK,  Screen SE,  Tabaska JE,  Cao Y;
XX
DR   WPI; 2004-180133/17.
XX
PT   New recombinant DNA construct, useful for improving plant tolerance to
PT   cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT   pests, for conferring increased resistance to plant disease, or for
PT   improving yield.
XX
PS   Claim 1; SEQ ID NO 70634; 15pp; English.
XX
CC   The invention describes a recombinant DNA construct comprising a
CC   polynucleotide consisting of a sequence encoding an amino acid sequence
CC   available in electronic form from the US patent office at
CC   ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC   of the invention are also useful in physical arrays of molecules and as
CC   plant breeding markers. The recombinant DNA construct is useful for
CC   improving plant tolerance to cold, heat, drought, herbicides, extreme
CC   osmotic conditions, pathogens or pests, for manipulating growth rate in
CC   plant cells by modification of the cell cycle pathway, for conferring
CC   increased resistance to plant disease, for producing galactomannan,
CC   lignin or plant growth regulators, for increasing the rate of homologous
CC   recombination in plants, for improving yield by modification of
CC   photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC   or by providing improved plant growth and development under at least one
CC   stress condition or for modifying seed oil or protein yield and/or
CC   content. This is the amino acid sequence of a plant full length insert
CC   polypeptide that can be used in the recombinant DNA construct of the
CC   invention.
XX
SQ   Sequence 382 AA;

  Query Match             90.2%;  Score 110;  DB 8;  Length 382;
  Best Local Similarity   95.7%;  Pred. No. 3.4e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLWT 23
              ||||||||||||||||||| |||
Db        109 LVAEVSEFLFGSGFAIAEGDLWT 131


RESULT 7
AEB16919
ID   AEB16919 standard; protein; 545 AA.
XX
AC   AEB16919;
XX
DT   08-SEP-2005  (first entry)
XX
DE   Barley cytochrome P450 97C (CYP97C) protein, SEQ ID NO: 17.
XX
KW   Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C;
KW   cytochrome P450 97C.
XX
OS   Hordeum vulgare; subsp. vulgare.
OS   Hordeum vulgare; subsp. spontaneum.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 529
FT                   /note= "Encoded by AGN"
XX
PN   US2005150002-A1.
XX
PD   07-JUL-2005.
XX
PF   02-JAN-2004; 2004US-00751235.
XX
PR   02-JAN-2004; 2004US-00751235.
XX
PA   (DELL/) DELLAPENNA D.
PA   (TIAN/) TIAN L.
PA   (KIMJ/) KIM J.
XX
PI   Dellapenna D,  Tian L,  Kim J;
XX
DR   WPI; 2005-487984/49.
DR   N-PSDB; AEB16925.
DR   GENBANK; BM816653, BU987393, CA023004.
DR   DDBJ; AV835803.
XX
PT   New expression vector comprising a nucleic acid sequence encoding a
PT   polypeptide having monooxygenase P450 activity, useful in altering the
PT   carotenoid production in a plant for enhancing production of specific
PT   carotenoid compounds.
XX
PS   Claim 9; SEQ ID NO 17; 135pp; English.
XX
CC   The present invention relates to genes, proteins and methods comprising
CC   carotenoid monooxygenases in the cytochrome P450 family. The invention
CC   also relates to altering carotenoid ratios in plants and microorganisms
CC   using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC   invention is useful in altering the carotenoid production in a plant for
CC   enhancing production of specific carotenoid compounds that are potent
CC   antioxidants. The present sequence is barley cytochrome P450
CC   monooxygenase (CYP97C) protein.
XX
SQ   Sequence 545 AA;
```

```
  Query Match             90.2%;  Score 110;  DB 9;  Length 545;
  Best Local Similarity   95.7%;  Pred. No. 5.1e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLWT 23
              ||||||||||||||||||||| |||
Db        136 LVAEVSEFLFGSGFAIAEGALWT 158


RESULT 8
AEB16918
ID   AEB16918 standard; protein; 561 AA.
XX
AC   AEB16918;
XX
DT   08-SEP-2005  (first entry)
XX
DE   Rice cytochrome P450 97C2 (CYP97C2) protein, SEQ ID NO: 16.
XX
KW   Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C2;
KW   cytochrome P450 97C2.
XX
OS   Oryza sativa; japonica cultivar-group.
XX
PN   US2005150002-A1.
XX
PD   07-JUL-2005.
XX
PF   02-JAN-2004; 2004US-00751235.
XX
PR   02-JAN-2004; 2004US-00751235.
XX
PA   (DELL/) DELLAPENNA D.
PA   (TIAN/) TIAN L.
PA   (KIMJ/) KIM J.
XX
PI   Dellapenna D,  Tian L,  Kim J;
XX
DR   WPI; 2005-487984/49.
DR   N-PSDB; AEB16924.
DR   GENBANK; AAK20054.
XX
PT   New expression vector comprising a nucleic acid sequence encoding a
PT   polypeptide having monooxygenase P450 activity, useful in altering the
PT   carotenoid production in a plant for enhancing production of specific
PT   carotenoid compounds.
XX
PS   Claim 9; SEQ ID NO 16; 135pp; English.
XX
CC   The present invention relates to genes, proteins and methods comprising
CC   carotenoid monooxygenases in the cytochrome P450 family. The invention
CC   also relates to altering carotenoid ratios in plants and microorganisms
CC   using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC   invention is useful in altering the carotenoid production in a plant for
CC   enhancing production of specific carotenoid compounds that are potent
CC   antioxidants. The present sequence is rice cytochrome P450 monooxygenase
CC   (CYP97C1) protein.
XX
SQ   Sequence 561 AA;

  Query Match             90.2%;  Score 110;  DB 9;  Length 561;
  Best Local Similarity   95.7%;  Pred. No. 5.3e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLWT 23
              ||||||||||||||||||||| |||
Db        152 LVAEVSEFLFGSGFAIAEGALWT 174


RESULT 9
AEB16954
ID   AEB16954 standard; protein; 588 AA.
XX
AC   AEB16954;
XX
DT   08-SEP-2005  (first entry)
XX
DE   Rice cytochrome P450 97B4 (CYP97B4) protein, SEQ ID NO: 52.
XX
KW   Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97B4;
KW   cytochrome P450 97B.
XX
OS   Oryza sativa; japonica cultivar-group.
XX
PN   US2005150002-A1.
XX
PD   07-JUL-2005.
XX
PF   02-JAN-2004; 2004US-00751235.
XX
PR   02-JAN-2004; 2004US-00751235.
XX
PA   (DELL/) DELLAPENNA D.
PA   (TIAN/) TIAN L.
PA   (KIMJ/) KIM J.
XX
PI   Dellapenna D,  Tian L,  Kim J;
XX
DR   WPI; 2005-487984/49.
DR   EMBL; E017117.
XX
PT   New expression vector comprising a nucleic acid sequence encoding a
PT   polypeptide having monooxygenase P450 activity, useful in altering the
PT   carotenoid production in a plant for enhancing production of specific
PT   carotenoid compounds.
XX
PS   Claim 9; SEQ ID NO 52; 135pp; English.
XX
CC   The present invention relates to genes, proteins and methods comprising
CC   carotenoid monooxygenases in the cytochrome P450 family. The invention
CC   also relates to altering carotenoid ratios in plants and microorganisms
CC   using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC   invention is useful in altering the carotenoid production in a plant for
CC   enhancing production of specific carotenoid compounds that are potent
CC   antioxidants. The present sequence is rice cytochrome P450 monooxygenase
CC   (CYP97B4) protein.
XX
SQ   Sequence 588 AA;

  Query Match             90.2%;  Score 110;  DB 9;  Length 588;
  Best Local Similarity   95.7%;  Pred. No. 5.6e-09;
  Matches   22;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLWT 23
              ||||||||||||||||||||| |||
Db        154 LVAEVSEFLFGSGFAIAEGALWT 176


RESULT 10
ABU25025
ID   ABU25025 standard; protein; 167 AA.
XX
AC   ABU25025;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Protein encoded by Prokaryotic essential gene #10552.
XX
KW   Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS   Clostridium difficile.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
```

```
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,  Zamudio C,  Malone C,  Haselbeck R,  Ohlsen KL,  Zyskind JW;
PI   Wall D,  Trawick JD,  Carr GJ,  Yamamoto R,  Forsyth RA,  Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   N-PSDB; ACA28895.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 25; SEQ ID NO 52949; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC   the target prokaryotic essential genes. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 167 AA;

  Query Match             48.4%;  Score 59;  DB 6;  Length 167;
  Best Local Similarity   50.0%;  Pred. No. 0.3;
  Matches   10;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGP 20
              :: ||::   | ||:|||||
Db        113 IIDEVTDIFIGMGFSIAEGP 132


RESULT 11
ABU25075
ID   ABU25075 standard; protein; 342 AA.
XX
AC   ABU25075;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Protein encoded by Prokaryotic essential gene #10602.
XX
KW   Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS   Clostridium difficile.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,  Zamudio C,  Malone C,  Haselbeck R,  Ohlsen KL,  Zyskind JW;
PI   Wall D,  Trawick JD,  Carr GJ,  Yamamoto R,  Forsyth RA,  Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   N-PSDB; ACA28945.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 25; SEQ ID NO 52999; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC   the target prokaryotic essential genes. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 342 AA;

  Query Match             48.4%;  Score 59;  DB 6;  Length 342;
  Best Local Similarity   50.0%;  Pred. No. 0.7;
  Matches   10;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGP 20
              :: ||::   | ||:|||||
Db        116 IIDEVTDIFIGMGFSIAEGP 135
```

```
RESULT 12
AEB16936
ID   AEB16936 standard; protein; 632 AA.
XX
AC   AEB16936;
XX
DT   08-SEP-2005  (first entry)
XX
DE   Rice cytochrome P450 97A (CYP97A) protein, SEQ ID NO: 34.
XX
KW   Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97A;
KW   cytochrome P450 97A.
XX
OS   Oryza sativa; japonica cultivar-group.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 93. .94
FT                   /note= "Encoded by bases of positions 277 to 379"
FT   Misc-difference 164. .165
FT                   /note= "Encoded by bases of positions 586 to 691"
FT   Misc-difference 188. .189
FT                   /note= "Encoded by bases of positions 758 to 850"
FT   Misc-difference 225. .226
FT                   /note= "Encoded by bases of positions 956 to 1042"
FT   Misc-difference 289. .290
FT                   /note= "Encoded by bases of positions 1229 to 1336"
FT   Misc-difference 343. .344
FT                   /note= "Encoded by bases of positions 1493 to 1584"
FT   Misc-difference 376. .377
FT                   /note= "Encoded by bases of positions 1678 to 1763"
FT   Misc-difference 409. .410
FT                   /note= "Encoded by bases of positions 1857 to 2009"
FT   Misc-difference 420. .421
FT                   /note= "Encoded by bases of positions 2037 to 2139"
FT   Misc-difference 447. .448
FT                   /note= "Encoded by bases of positions 2015 to 2301"
FT   Misc-difference 475. .476
FT                   /note= "Encoded by bases of positions 2380 to 3051"
FT   Misc-difference 518. .519
FT                   /note= "Encoded by bases of positions 3175 to 4418"
FT   Misc-difference 539. .540
FT                   /note= "Encoded by bases of positions 4476 to 4715"
FT   Misc-difference 562. .563
FT                   /note= "Encoded by bases of positions 4776 to 4856"
XX
PN   US2005150002-A1.
XX
PD   07-JUL-2005.
XX
PF   02-JAN-2004; 2004US-00751235.
XX
PR   02-JAN-2004; 2004US-00751235.
XX
PA   (DELL/) DELLAPENNA D.
PA   (TIAN/) TIAN L.
PA   (KIMJ/) KIM J.
XX
PI   Dellapenna D,  Tian L,  Kim J;
XX
DR   WPI; 2005-487984/49.
DR   N-PSDB; AEB16944, AEB16945.
DR   DDBJ; AP004028.
XX
PT   New expression vector comprising a nucleic acid sequence encoding a
PT   polypeptide having monooxygenase P450 activity, useful in altering the
PT   carotenoid production in a plant for enhancing production of specific
PT   carotenoid compounds.
XX
PS   Claim 9; SEQ ID NO 34; 135pp; English.
XX
CC   The present invention relates to genes, proteins and methods comprising
CC   carotenoid monooxygenases in the cytochrome P450 family. The invention
CC   also relates to altering carotenoid ratios in plants and microorganisms
CC   using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC   invention is useful in altering the carotenoid production in a plant for
CC   enhancing production of specific carotenoid compounds that are potent
CC   antioxidants. The present sequence is rice cytochrome P450 monooxygenase
CC   (CYP97A) protein.
XX
SQ   Sequence 632 AA;

  Query Match             47.5%;  Score 58;  DB 9;  Length 632;
  Best Local Similarity   40.9%;  Pred. No. 2.1;
  Matches    9;  Conservative    7;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLW 22
              ::||: ||: |:|   |:| :|
Db        190 ILAEILEFVMGTGLIPADGEIW 211


RESULT 13
ADX67984
ID   ADX67984 standard; protein; 662 AA.
XX
AC   ADX67984;
XX
DT   21-APR-2005  (first entry)
XX
DE   Plant full length insert polypeptide seqid 38827.
XX
KW   plant protectant; plant growth regulant; gene therapy; plant;
KW   recombinant DNA construct; physical array; plant breeding marker;
KW   cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW   extreme osmotic condition; pathogen tolerance; pest tolerance;
KW   growth rate; cell cycle pathway; disease resistance;
KW   galactomannan production; lignin production; plant growth regulator;
KW   yield; plant growth; plant development; seed oil; protein yield;
KW   protein content.
XX
OS   Unidentified.
XX
PN   US2004034888-A1.
XX
PD   19-FEB-2004.
XX
PF   28-APR-2003; 2003US-00425114.
XX
PR   06-MAY-1999;   99US-00304517.
PR   05-NOV-2001; 2001US-00985678.
XX
PA   (LIUJ/) LIU J.
PA   (ZHOU/) ZHOU Y.
PA   (KOVA/) KOVALIC D K.
PA   (SCRE/) SCREEN S E.
PA   (TABA/) TABASKA J E.
PA   (CAOY/) CAO Y.
XX
PI   Liu J,  Zhou Y,  Kovalic DK,  Screen SE,  Tabaska JE,  Cao Y;
XX
DR   WPI; 2004-180133/17.
XX
PT   New recombinant DNA construct, useful for improving plant tolerance to
PT   cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT   pests, for conferring increased resistance to plant disease, or for
PT   improving yield.
XX
PS   Claim 1; SEQ ID NO 38827; 15pp; English.
XX
CC   The invention describes a recombinant DNA construct comprising a
CC   polynucleotide consisting of a sequence encoding an amino acid sequence
CC   available in electronic form from the US patent office at
CC   ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC   of the invention are also useful in physical arrays of molecules and as
CC   plant breeding markers. The recombinant DNA construct is useful for
```

```
CC   improving plant tolerance to cold, heat, drought, herbicides, extreme
CC   osmotic conditions, pathogens or pests, for manipulating growth rate in
CC   plant cells by modification of the cell cycle pathway, for conferring
CC   increased resistance to plant disease, for producing galactomannan,
CC   lignin or plant growth regulators, for increasing the rate of homologous
CC   recombination in plants, for improving yield by modification of
CC   photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC   or by providing improved plant growth and development under at least one
CC   stress condition or for modifying seed oil or protein yield and/or
CC   content. This is the amino acid segeunce of a plant full length insert
CC   polypeptide that can be used in the recombinant DNA construct of the
CC   invention.
XX
SQ   Sequence 662 AA;

  Query Match             47.5%;  Score 58;  DB 8;  Length 662;
  Best Local Similarity   40.9%;  Pred. No. 2.2;
  Matches    9;  Conservative    7;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLW 22
              ::||: ||: |:|   |:| :|
Db        225 ILAEILEFVMGTGLIPADGEIW 246


RESULT 14
AEB16939
ID   AEB16939 standard; protein; 342 AA.
XX
AC   AEB16939;
XX
DT   08-SEP-2005  (first entry)
XX
DE   Wheat cytochrome P450 97A (CYP97A) protein, SEQ ID NO: 37.
XX
KW   Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97A;
KW   cytochrome P450 97A.
XX
OS   Triticum aestivum.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 197
FT                   /note= "Encoded by CAGAG"
XX
PN   US2005150002-A1.
XX
PD   07-JUL-2005.
XX
PF   02-JAN-2004; 2004US-00751235.
XX
PR   02-JAN-2004; 2004US-00751235.
XX
PA   (DELL/) DELLAPENNA D.
PA   (TIAN/) TIAN L.
PA   (KIMJ/) KIM J.
XX
PI   Dellapenna D,  Tian L,  Kim J;
XX
DR   WPI; 2005-487984/49.
DR   N-PSDB; AEB16948.
DR   DDBJ; BJ234910.
DR   GENBANK; CA736787, CA736801.
DR   DDBJ; BJ238659, BJ233019.
DR   GENBANK; CD882035.
XX
PT   New expression vector comprising a nucleic acid sequence encoding a
PT   polypeptide having monooxygenase P450 activity, useful in altering the
PT   carotenoid production in a plant for enhancing production of specific
PT   carotenoid compounds.
XX
PS   Claim 9; SEQ ID NO 37; 135pp; English.
XX
CC   The present invention relates to genes, proteins and methods comprising
CC   carotenoid monooxygenases in the cytochrome P450 family. The invention
CC   also relates to altering carotenoid ratios in plants and microorganisms
CC   using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC   invention is useful in altering the carotenoid production in a plant for
CC   enhancing production of specific carotenoid compounds that are potent
CC   antioxidants. The present sequence is wheat cytochrome P450 monooxygenase
CC   (CYP97A) protein.
XX
SQ   Sequence 342 AA;

  Query Match             46.7%;  Score 57;  DB 9;  Length 342;
  Best Local Similarity   40.9%;  Pred. No. 1.5;
  Matches    9;  Conservative    7;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLW 22
              ::||: ||: |:|   |:| :|
Db         67 ILAEILEFVMGTGLIPADGEVW 88


RESULT 15
AEB16937
ID   AEB16937 standard; protein; 508 AA.
XX
AC   AEB16937;
XX
DT   08-SEP-2005  (first entry)
XX
DE   Barley cytochrome P450 97A (CYP97A) protein, SEQ ID NO: 35.
XX
KW   Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97A;
KW   cytochrome P450 97A.
XX
OS   Hordeum vulgare; subsp. vulgare.
OS   Hordeum vulgare; subsp. spontaneum.
XX
PN   US2005150002-A1.
XX
PD   07-JUL-2005.
XX
PF   02-JAN-2004; 2004US-00751235.
XX
PR   02-JAN-2004; 2004US-00751235.
XX
PA   (DELL/) DELLAPENNA D.
PA   (TIAN/) TIAN L.
PA   (KIMJ/) KIM J.
XX
PI   Dellapenna D,  Tian L,  Kim J;
XX
DR   WPI; 2005-487984/49.
DR   N-PSDB; AEB16946.
DR   DDBJ; AV939715, AV941342, AV939552, AV939556.
DR   GENBANK; CA004011.
DR   DDBJ; BJ480615, BJ485000, BJ448041, BJ455787, AV910152, AV938407.
DR   EMBL; AJ477620, AJ477618, AJ477619, AJ832622.
XX
PT   New expression vector comprising a nucleic acid sequence encoding a
PT   polypeptide having monooxygenase P450 activity, useful in altering the
PT   carotenoid production in a plant for enhancing production of specific
PT   carotenoid compounds.
XX
PS   Claim 9; SEQ ID NO 35; 135pp; English.
XX
CC   The present invention relates to genes, proteins and methods comprising
CC   carotenoid monooxygenases in the cytochrome P450 family. The invention
CC   also relates to altering carotenoid ratios in plants and microorganisms
CC   using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC   invention is useful in altering the carotenoid production in a plant for
CC   enhancing production of specific carotenoid compounds that are potent
CC   antioxidants. The present sequence is barley cytochrome P450
CC   monooxygenase (CYP97A) protein.
XX
SQ   Sequence 508 AA;
```

Query Match 46.7%; Score 57; DB 9; Length 508;
Best Local Similarity 40.9%; Pred. No. 2.4;
Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 LVAEVSEFLFGSGFAIAEGPLW 22
              ::||: ||: |:|   |:| :|
Db         66 ILAEILEFVMGTGLIPADGEVW 87
```

Search completed: March 7, 2006, 21:53:06
Job time : 240.8 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2006, 21:58:31 ; Search time 26.8 Seconds
(without alignments)
74.038 Million cell updates/sec

Title: US-10-751-235-10
Perfect score: 122
Sequence: 1 LVAEVSEFLFGSGFAIAEGPLWTA 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	52.5	43.0	94	2	US-09-605-703B-1572	Sequence 1572, Ap
2	49	40.2	194	2	US-09-270-767-59102	Sequence 59102, A
3	49	40.2	230	2	US-09-270-767-43713	Sequence 43713, A
4	48	39.3	316	2	US-09-902-540-10421	Sequence 10421, A
5	47	38.5	683	2	US-09-252-991A-28674	Sequence 28674, A
6	46.5	38.1	190	2	US-09-134-000C-6230	Sequence 6230, Ap
7	46	37.7	55	2	US-09-513-999C-4532	Sequence 4532, Ap
8	46	37.7	80	2	US-09-621-976-5244	Sequence 5244, Ap
9	46	37.7	150	2	US-09-270-767-60566	Sequence 60566, A
10	46	37.7	250	2	US-09-489-039A-10680	Sequence 10680, A
11	45	36.9	102	2	US-09-882-434A-1	Sequence 1, Appli
12	45	36.9	110	2	US-09-376-330-24	Sequence 24, Appl
13	45	36.9	131	2	US-09-513-999C-8015	Sequence 8015, Ap
14	45	36.9	166	2	US-09-621-976-4638	Sequence 4638, Ap
15	45	36.9	185	2	US-09-529-157-6	Sequence 6, Appli
16	45	36.9	185	2	US-09-529-157-7	Sequence 7, Appli
17	45	36.9	204	2	US-09-949-016-9489	Sequence 9489, Ap
18	45	36.9	205	2	US-09-673-395A-205	Sequence 205, App
19	45	36.9	218	2	US-09-543-681A-4284	Sequence 4284, Ap
20	45	36.9	329	2	US-10-152-886-51	Sequence 51, Appl
21	45	36.9	522	1	US-08-305-505-2	Sequence 2, Appli
22	45	36.9	561	2	US-09-489-039A-10642	Sequence 10642, A
23	45	36.9	862	2	US-09-902-540-11888	Sequence 11888, A
24	44	36.1	194	2	US-09-252-991A-31239	Sequence 31239, A
25	44	36.1	519	2	US-09-248-796A-19188	Sequence 19188, A
26	44	36.1	865	1	US-07-803-633A-13	Sequence 13, Appl
27	43	35.2	82	2	US-09-252-991A-23597	Sequence 23597, A
28	43	35.2	188	2	US-09-134-000C-3652	Sequence 3652, Ap
29	43	35.2	297	2	US-09-902-540-14556	Sequence 14556, A
30	43	35.2	310	2	US-09-252-991A-25022	Sequence 25022, A
31	43	35.2	392	2	US-09-902-540-10171	Sequence 10171, A
32	43	35.2	411	2	US-09-252-991A-31301	Sequence 31301, A
33	43	35.2	642	2	US-10-152-886-9	Sequence 9, Appli
34	43	35.2	871	2	US-09-548-473B-7	Sequence 7, Appli
35	43	35.2	1351	2	US-09-548-473B-1	Sequence 1, Appli
36	43	35.2	1548	2	US-09-252-991A-22301	Sequence 22301, A
37	43	35.2	1610	2	US-09-548-473B-4	Sequence 4, Appli
38	43	35.2	1665	2	US-09-858-664A-2	Sequence 2, Appli
39	43	35.2	1665	2	US-10-274-978-2	Sequence 2, Appli
40	43	35.2	1665	2	US-10-697-263-2	Sequence 2, Appli
41	43	35.2	1722	2	US-09-538-092-1033	Sequence 1033, Ap
42	43	35.2	2596	2	US-09-548-473B-6	Sequence 6, Appli
43	43	35.2	3170	1	US-07-642-734C-5	Sequence 5, Appli
44	43	35.2	3170	2	US-08-439-009A-5	Sequence 5, Appli
45	42.5	34.8	404	2	US-09-560-761B-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-09-605-703B-1572
; Sequence 1572, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; TITLE OF INVENTION: PROTEINS
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 1572
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-605-703B-1572

Query Match 43.0%; Score 52.5; DB 2; Length 94;
Best Local Similarity 57.7%; Pred. No. 0.41;
Matches 15; Conservative 1; Mismatches 7; Indels 3; Gaps 1;

```
Qy          1 LVAEVSEFL---FGSGFAIAEGPLWT 23
              |:| ||  |   |||| |||   |||
Db          6 LLAMVSTALRIRFGSGVAIAATVLWT 31
```

RESULT 2
US-09-270-767-59102
; Sequence 59102, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59102
; LENGTH: 194
; TYPE: PRT

; ORGANISM: Drosophila melanogaster
US-09-270-767-59102

Query Match 40.2%; Score 49; DB 2; Length 194;
Best Local Similarity 75.0%; Pred. No. 3.7;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 8 FLFGSGFAIAEG 19
|| | |||:|||
Db 128 FLLGGGFALAEG 139

RESULT 3
US-09-270-767-43713
; Sequence 43713, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43713
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-43713

Query Match 40.2%; Score 49; DB 2; Length 230;
Best Local Similarity 75.0%; Pred. No. 4.6;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 8 FLFGSGFAIAEG 19
|| | |||:|||
Db 164 FLLGGGFALAEG 175

RESULT 4
US-09-902-540-10421
; Sequence 10421, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10421
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10421

Query Match 39.3%; Score 48; DB 2; Length 316;
Best Local Similarity 45.5%; Pred. No. 9.8;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 1 LVAEVSEFLFGSGFAIAEGPLW 22
|| | :| :|||: |||
Db 261 LVQGVESGMFQAGFALNPPPLW 282

RESULT 5
US-09-252-991A-28674
; Sequence 28674, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28674
; LENGTH: 683
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28674

Query Match 38.5%; Score 47; DB 2; Length 683;
Best Local Similarity 56.2%; Pred. No. 36;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 9 LFGSGFAIAEGPLWTA 24
|||| :::||| | |
Db 440 LFGSFYSVAEGLEWLA 455

RESULT 6
US-09-134-000C-6230
; Sequence 6230, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6230
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6230

Query Match 38.1%; Score 46.5; DB 2; Length 190;
Best Local Similarity 50.0%; Pred. No. 9.4;
Matches 12; Conservative 2; Mismatches 5; Indels 5; Gaps 1;

Qy 1 LVAEVSEFLFG-----SGFAIAEG 19
|| :|| ||| ||| |: |
Db 91 LVFSISELLFGLAQAKSGFYISRG 114

RESULT 7
US-09-513-999C-4532
; Sequence 4532, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26

```
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 4532
;   LENGTH: 55
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIGNAL
;   LOCATION: -23..-1
;   OTHER INFORMATION: score 9.4
;   OTHER INFORMATION: seq SLLFICFFGESFC/IC
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 9
;   OTHER INFORMATION: Xaa= *  or Lys or Leu or Met
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 11
;   OTHER INFORMATION: Xaa= *  or Glu or Gly
US-09-513-999C-4532

  Query Match             37.7%;  Score 46;  DB 2;  Length 55;
  Best Local Similarity   50.0%;  Pred. No. 2.6;
  Matches    7;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         10 FGSGFAIAEGPLWT 23
              ||  | | :| :||
Db         18 FGESFCICDGTVWT 31


RESULT 8
US-09-621-976-5244
; Sequence 5244, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 5244
;   LENGTH: 80
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SIGNAL
;   LOCATION: -23..-1
US-09-621-976-5244

  Query Match             37.7%;  Score 46;  DB 2;  Length 80;
  Best Local Similarity   50.0%;  Pred. No. 4;
  Matches    7;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         10 FGSGFAIAEGPLWT 23
              ||  | | :| :||
Db         18 FGESFCICDGTVWT 31


RESULT 9
US-09-270-767-60566
; Sequence 60566, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60566
;   LENGTH: 150
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
;   FEATURE:
;   OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-60566

  Query Match             37.7%;  Score 46;  DB 2;  Length 150;
  Best Local Similarity   52.6%;  Pred. No. 8.6;
  Matches   10;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          6 SEPLFGSGFAIAEGPLWTA 24
              |:|:    |||:||||  |
Db         36 SKFVSAQRFAISEGPLCCA 54


RESULT 10
US-09-489-039A-10680
; Sequence 10680, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 10680
;   LENGTH: 250
;   TYPE: PRT
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10680

  Query Match             37.7%;  Score 46;  DB 2;  Length 250;
  Best Local Similarity   72.7%;  Pred. No. 16;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         12 SGFAIAEGPLW 22
              |||  ||||:|
Db         93 SGFRWAEGPVW 103


RESULT 11
US-09-882-434A-1
; Sequence 1, Application US/09882434A
; Patent No. 6909032
; GENERAL INFORMATION:
;  APPLICANT: Manners, John M.
;  APPLICANT:  Marcus, John Paul
;  APPLICANT:  Goulter, Kenneth C.
;  APPLICANT:  Green, Jodie Lyn
;  TITLE OF INVENTION: ANTI-MICROBIAL PROTEIN
;  FILE REFERENCE: CULLN18.1CP1C1
;  CURRENT APPLICATION NUMBER: US/09/882,434A
;  CURRENT FILING DATE:  2001-06-15
;  PRIOR APPLICATION NUMBER: 09/364395
;  PRIOR FILING DATE: 1999-07-30
;  PRIOR APPLICATION NUMBER: 09/117615
;  PRIOR FILING DATE: 1998-11-09
;  PRIOR APPLICATION NUMBER: PCT/AU97/00052
;  PRIOR FILING DATE: 1997-01-31
;  PRIOR APPLICATION NUMBER: AU PN 7802
;  PRIOR FILING DATE: 1996-01-31
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 102
;   TYPE: PRT
```

```
;   ORGANISM: Macadamia integrifolia
US-09-882-434A-1

  Query Match             36.9%;  Score 45;  DB 2;  Length 102;
  Best Local Similarity   45.0%;  Pred. No. 7.9;
  Matches    9;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGP 20
              |:|  || : || | :  ||
Db         16 LIAMASEMVNGSAFTVWSGP 35


RESULT 12
US-09-376-330-24
; Sequence 24, Application US/09376330
; Patent No. 6399321
; GENERAL INFORMATION:
;  APPLICANT: Tessier, Daniel C.
;  APPLICANT:  Dignard, Daniel
;  APPLICANT:  Bergeron, John J.M.
;  APPLICANT:  Thomas, David Y.
;  TITLE OF INVENTION: Method for screening for
;  TITLE OF INVENTION:  UDP-glucose:glycoprotein glucosyltransferase (UGGT) activity
;  TITLE OF INVENTION:  and nucleic acid encoding for UGGT
;  FILE REFERENCE: 2139-9"US"
;  CURRENT APPLICATION NUMBER: US/09/376,330
;  CURRENT FILING DATE:  1999-08-18
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 24
;   LENGTH: 110
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: GALT-S. typ
US-09-376-330-24

  Query Match             36.9%;  Score 45;  DB 2;  Length 110;
  Best Local Similarity   53.3%;  Pred. No. 8.6;
  Matches    8;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         10 FGSGFAIAEGPLWTA 24
              | :|| :   |||||
Db         62 FNAGFILIXIPLWTA 76


RESULT 13
US-09-513-999C-8015
; Sequence 8015, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 8015
;   LENGTH: 131
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-513-999C-8015

  Query Match             36.9%;  Score 45;  DB 2;  Length 131;
  Best Local Similarity   36.4%;  Pred. No. 11;
  Matches    8;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLW 22
              | |: |   ||: | ::  |:|
Db         24 LSAKSSALFFGNAFIVSAIPIW 45


RESULT 14
US-09-621-976-4638
; Sequence 4638, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 4638
;   LENGTH: 166
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-621-976-4638

  Query Match             36.9%;  Score 45;  DB 2;  Length 166;
  Best Local Similarity   36.4%;  Pred. No. 14;
  Matches    8;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLW 22
              | |: |   ||: | ::  |:|
Db         24 LSAKSSALFFGNAFIVSAIPIW 45


RESULT 15
US-09-529-157-6
; Sequence 6, Application US/09529157
; Patent No. 6500939
; GENERAL INFORMATION:
;  APPLICANT: Kato, Seishi
;  APPLICANT:  Sekine, Shingo
;  TITLE OF INVENTION: cDNAs Coding For Human Proteins Having Transmembrane
;  TITLE OF INVENTION:  Domains
;  FILE REFERENCE: GIN-6711CPUS
;  CURRENT APPLICATION NUMBER: US/09/529,157
;  CURRENT FILING DATE:  2000-08-21
;  PRIOR APPLICATION NUMBER: PCT/JP98/04447
;  PRIOR FILING DATE: 1998-10-02
;  PRIOR APPLICATION NUMBER: JP 9-276270
;  PRIOR FILING DATE: 1997-10-08
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 185
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-529-157-6

  Query Match             36.9%;  Score 45;  DB 2;  Length 185;
  Best Local Similarity   36.4%;  Pred. No. 16;
  Matches    8;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAEGPLW 22
              | |: |   ||: | ::  |:|
Db         24 LSAKSSALFFGNAFIVSAIPIW 45


Search completed: March  7, 2006, 22:00:14
Job time : 27.8 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2006, 22:00:32 ; Search time 10.4 Seconds
(without alignments)
46.155 Million cell updates/sec

Title: US-10-751-235-10
Perfect score: 122
Sequence: 1 LVAEVSEFLFGSGFAIAEGPLWTA 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 135346 seqs, 20000420 residues

Total number of hits satisfying chosen parameters: 135346

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	48.5	39.8	565	7	US-11-087-099-12143	Sequence 12143, A
2	48.5	39.8	613	7	US-11-087-099-5132	Sequence 5132, Ap
3	48.5	39.8	654	7	US-11-087-099-10545	Sequence 10545, A
4	46	37.7	246	6	US-10-993-143-19	Sequence 19, Appl
5	46	37.7	346	7	US-11-098-686-10717	Sequence 10717, A
6	44	36.1	295	6	US-10-858-730-112	Sequence 112, App
7	44	36.1	364	6	US-10-467-657-2822	Sequence 2822, Ap
8	44	36.1	505	7	US-11-087-099-6346	Sequence 6346, Ap
9	43	35.2	301	7	US-11-152-569-9	Sequence 9, Appli
10	43	35.2	491	7	US-11-087-099-3329	Sequence 3329, Ap
11	43	35.2	576	7	US-11-087-099-10787	Sequence 10787, A
12	43	35.2	2630	7	US-11-186-731-2	Sequence 2, Appli
13	43	35.2	7968	7	US-11-186-731-5	Sequence 5, Appli
14	42.5	34.8	461	7	US-11-087-099-11993	Sequence 11993, A
15	42	34.4	444	7	US-11-072-512-2690	Sequence 2690, Ap
16	42	34.4	509	6	US-10-131-826A-108	Sequence 108, App
17	42	34.4	509	6	US-10-973-115B-108	Sequence 108, App
18	42	34.4	583	7	US-11-087-099-6900	Sequence 6900, Ap
19	42	34.4	893	7	US-11-230-145-4	Sequence 4, Appli
20	41.5	34.0	180	7	US-11-087-099-1577	Sequence 1577, Ap
21	41.5	34.0	180	7	US-11-087-099-10964	Sequence 10964, A
22	41.5	34.0	263	6	US-10-873-528-167	Sequence 167, App
23	41.5	34.0	524	7	US-11-054-281-34	Sequence 34, Appl
24	41.5	34.0	524	7	US-11-054-281-121	Sequence 121, App
25	41.5	34.0	541	7	US-11-098-686-11058	Sequence 11058, A
26	41	33.6	227	6	US-10-467-657-2594	Sequence 2594, Ap
27	41	33.6	261	7	US-11-087-099-5386	Sequence 5386, Ap
28	41	33.6	318	7	US-11-087-099-4152	Sequence 4152, Ap
29	41	33.6	388	6	US-10-995-561-887	Sequence 887, App
30	41	33.6	563	7	US-11-087-099-716	Sequence 716, App
31	41	33.6	640	7	US-11-087-099-6404	Sequence 6404, Ap
32	40.5	33.2	691	6	US-10-467-657-7170	Sequence 7170, Ap
33	40.5	33.2	801	6	US-10-793-626-2020	Sequence 2020, Ap
34	40	32.8	283	7	US-11-087-099-9654	Sequence 9654, Ap
35	40	32.8	312	7	US-11-087-099-5676	Sequence 5676, Ap
36	40	32.8	453	7	US-11-087-099-12171	Sequence 12171, A
37	40	32.8	610	7	US-11-069-642-22	Sequence 22, Appl
38	39.5	32.4	120	6	US-10-467-657-3686	Sequence 3686, Ap
39	39	32.0	110	7	US-11-112-240-28	Sequence 28, Appl
40	39	32.0	110	7	US-11-112-304A-28	Sequence 28, Appl
41	39	32.0	130	6	US-10-993-543-144	Sequence 144, App
42	39	32.0	221	6	US-10-714-887-132	Sequence 132, App
43	39	32.0	243	7	US-11-054-515-1927	Sequence 1927, Ap
44	39	32.0	245	7	US-11-054-515-1926	Sequence 1926, Ap
45	39	32.0	245	7	US-11-054-515-3241	Sequence 3241, Ap

ALIGNMENTS

RESULT 1
US-11-087-099-12143
; Sequence 12143, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 12143
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(565)
; OTHER INFORMATION: unsure at all Xaa locations
US-11-087-099-12143

Query Match 39.8%; Score 48.5; DB 7; Length 565;
Best Local Similarity 33.3%; Pred. No. 6.7;
Matches 10; Conservative 6; Mismatches 5; Indels 9; Gaps 1;

```
Qy          2 VAEVSEFLFGSGFAIA---------EGPLW 22
              ::  |:| ||||| ::         :|| |
Db        534 ISGASDFRFGSGFNMSGTVDYPQKNDGPNW 563
```

RESULT 2
US-11-087-099-5132
; Sequence 5132, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 5132
; LENGTH: 613
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(613)

; OTHER INFORMATION: unsure at all Xaa locations
US-11-087-099-5132

Query Match 39.8%; Score 48.5; DB 7; Length 613;
Best Local Similarity 33.3%; Pred. No. 7.3;
Matches 10; Conservative 6; Mismatches 5; Indels 9; Gaps 1;

Qy 2 VAEVSEFLFGSGFAIA---------EGPLW 22
:: |:| ||||| :: :|| |
Db 582 ISGASDFRFGSGFNMSGTVDYPQKNDGPNW 611

RESULT 3
US-11-087-099-10545
; Sequence 10545, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 10545
; LENGTH: 654
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(654)
; OTHER INFORMATION: unsure at all Xaa locations
US-11-087-099-10545

Query Match 39.8%; Score 48.5; DB 7; Length 654;
Best Local Similarity 33.3%; Pred. No. 7.8;
Matches 10; Conservative 6; Mismatches 5; Indels 9; Gaps 1;

Qy 2 VAEVSEFLFGSGFAIA---------EGPLW 22
:: |:| ||||| :: :|| |
Db 623 ISGASDFRFGSGFNMSGTVDYPQKNDGPNW 652

RESULT 4
US-10-993-143-19
; Sequence 19, Application US/10993143
; Publication No. US20060036374A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology
; APPLICANT: Debe, Derek A.
; APPLICANT: Goddard III, William A.
; TITLE OF INVENTION: METHOD FOR DETERMINING THREE-DIMENSIONAL PROTEIN STRUCTURE FROM
; TITLE OF INVENTION: PRIMARY PROTEIN SEQUENCE
; FILE REFERENCE: 54318.8001.US02
; CURRENT APPLICATION NUMBER: US/10/993,143
; CURRENT FILING DATE: 2004-11-18
; PRIOR APPLICATION NUMBER: 60/218,016
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 09/905,176
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 19
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Rhodopseudomonas viridis
US-10-993-143-19

Query Match 37.7%; Score 46; DB 6; Length 246;
Best Local Similarity 56.2%; Pred. No. 6.8;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 8 FLFGSGFAIAEGPLWT 23
|| |: || || ||
Db 229 FLTGAFGTIASGPFWT 244

RESULT 5
US-11-098-686-10717
; Sequence 10717, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10717
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10717

Query Match 37.7%; Score 46; DB 7; Length 346;
Best Local Similarity 50.0%; Pred. No. 9.7;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 LVAEVSEFLFGSGFAIAEGP 20
||: : || | |||:| |
Db 207 LVSIIQGFLCGIGFAVANVP 226

RESULT 6
US-10-858-730-112
; Sequence 112, Application US/10858730
; Publication No. US20050255568A1
; GENERAL INFORMATION:
; APPLICANT: Bailey, Richard B.
; APPLICANT: Blomquist, Paul
; APPLICANT: Doten, Reed
; APPLICANT: Driggers, Edward M.
; APPLICANT: Madden, Kevin T.
; APPLICANT: O'Leary, Jessica
; APPLICANT: O'Toole, George
; APPLICANT: Trueheart, Joshua
; APPLICANT: Walbridge, Michael J.
; APPLICANT: Yorgey, Peter S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
; TITLE OF INVENTION: PRODUCTION
; FILE REFERENCE: 14184-030001
; CURRENT APPLICATION NUMBER: US/10/858,730
; CURRENT FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/475,000
; PRIOR FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: US 60/551,860
; PRIOR FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 112
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Streptomyces coelicolor
US-10-858-730-112

Query Match 36.1%; Score 44; DB 6; Length 295;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 11 GSGFAIAEGPLWTA 24
|: ||:| | :| |

```
Db        149 GAAFALAAGAMWAA 162


RESULT 7
US-10-467-657-2822
; Sequence 2822, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
;  APPLICANT: CHIRON SpA
;  APPLICANT: FONTANA Maria Rita
;  APPLICANT: PIZZA Mariagrazia
;  APPLICANT: MASIGNANI Vega
;  APPLICANT: MONACI Elisabetta
;  TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
;  FILE REFERENCE:
;  CURRENT APPLICATION NUMBER: US/10/467,657
;  CURRENT FILING DATE:  2003-08-11
;  PRIOR APPLICATION NUMBER: GB-0103424.8
;  PRIOR FILING DATE: 2001-02-12
;  NUMBER OF SEQ ID NOS: 9218
;  SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 2822
;    LENGTH: 364
;    TYPE: PRT
;    ORGANISM: Neisseria gonorrhoeae
US-10-467-657-2822

  Query Match             36.1%;  Score 44;  DB 6;  Length 364;
  Best Local Similarity   50.0%;  Pred. No. 21;
  Matches    8;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          5 VSEFLFGSGFAIAEGP 20
              | |   | || :|:||
Db        151 VVELFHGMGFEVADGP 166


RESULT 8
US-11-087-099-6346
; Sequence 6346, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
;  APPLICANT: Abad, Mark S. et al.
;  TITLE OF INVENTION: Genes and Uses for Plant Improvement
;  FILE REFERENCE: 38-21(53450)B EP
;  CURRENT APPLICATION NUMBER: US/11/087,099
;  CURRENT FILING DATE:  2005-03-22
;  NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 6346
;    LENGTH: 505
;    TYPE: PRT
;    ORGANISM: Neurospora crassa
US-11-087-099-6346

  Query Match             36.1%;  Score 44;  DB 7;  Length 505;
  Best Local Similarity   57.1%;  Pred. No. 29;
  Matches    8;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          9 LFGSGFAIAEGPLW 22
              |||:|   |||  |
Db        120 LFGNGIINAEGDFW 133


RESULT 9
US-11-152-569-9
; Sequence 9, Application US/11152569
; Publication No. US20060005278A1
; GENERAL INFORMATION:
;  APPLICANT: Ecker, Joseph R.
;  APPLICANT:  Nehring, Ramlah
;  APPLICANT:  McGrath, Robert B.
;  TITLE OF INVENTION: ETHYLENE INSENSITIVE PLANTS
;  FILE REFERENCE: 532792001210
;  CURRENT APPLICATION NUMBER: US/11/152,569
;  CURRENT FILING DATE:  2005-06-13
;  PRIOR APPLICATION NUMBER: 10/144,156
;  PRIOR FILING DATE: 2002-10-05
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
;    LENGTH: 301
;    TYPE: PRT
;    ORGANISM: Homo sapiens
US-11-152-569-9

  Query Match             35.2%;  Score 43;  DB 7;  Length 301;
  Best Local Similarity   47.1%;  Pred. No. 24;
  Matches    8;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 AEVSEFLFGSGFAIAEG 19
              |::|   ||||| : :|
Db         94 ADISSKDFGSGFPVKDG 110


RESULT 10
US-11-087-099-3329
; Sequence 3329, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
;  APPLICANT: Abad, Mark S. et al.
;  TITLE OF INVENTION: Genes and Uses for Plant Improvement
;  FILE REFERENCE: 38-21(53450)B EP
;  CURRENT APPLICATION NUMBER: US/11/087,099
;  CURRENT FILING DATE:  2005-03-22
;  NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 3329
;    LENGTH: 491
;    TYPE: PRT
;    ORGANISM: Pseudomonas aeruginosa PAO1
US-11-087-099-3329

  Query Match             35.2%;  Score 43;  DB 7;  Length 491;
  Best Local Similarity   53.3%;  Pred. No. 40;
  Matches    8;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 EVSEFLFGSGFAIAE 18
              ||   :||:|||  |
Db        334 EVDAIIFGTGFAATE 348


RESULT 11
US-11-087-099-10787
; Sequence 10787, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
;  APPLICANT: Abad, Mark S. et al.
;  TITLE OF INVENTION: Genes and Uses for Plant Improvement
;  FILE REFERENCE: 38-21(53450)B EP
;  CURRENT APPLICATION NUMBER: US/11/087,099
;  CURRENT FILING DATE:  2005-03-22
;  NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 10787
;    LENGTH: 576
;    TYPE: PRT
;    ORGANISM: Solanum tuberosum
US-11-087-099-10787

  Query Match             35.2%;  Score 43;  DB 7;  Length 576;
  Best Local Similarity   44.4%;  Pred. No. 48;
  Matches    8;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 LVAEVSEFLFGSGFAIAE 18
              ||| . :|:||:| |: :
Db        408 LVAGTVDFIFGNGAAVLQ 425


RESULT 12
```

```
US-11-186-731-2
; Sequence 2, Application US/11186731
; Publication No. US20050255521A1
; GENERAL INFORMATION:
;  APPLICANT: Kapeller-Libermann, Rosana
;  APPLICANT:  Acton, Susan L.
;  TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
;  TITLE OF INVENTION:  Members and Uses Therefor
;  FILE REFERENCE: MPI2001-047P1RCP1(M)
;  CURRENT APPLICATION NUMBER: US/11/186,731
;  CURRENT FILING DATE:  2005-07-21
;  PRIOR APPLICATION NUMBER: US/10/077,130
;  PRIOR FILING DATE: 2002-02-15
;  PRIOR APPLICATION NUMBER: 60/269201
;  PRIOR FILING DATE: 2001-02-15
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 2630
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-186-731-2

  Query Match             35.2%;  Score 43;  DB 7;  Length 2630;
  Best Local Similarity   39.1%;  Pred. No. 2.3e+02;
  Matches    9;  Conservative    4;  Mismatches    8;  Indels    2;  Gaps    1;

Qy          2 VAEVSEFLFG--SGFAIAEGPLW 22
              ||    | ||  :|  : :||:|
Db       1850 VAGYGTFAFGGDAGGMLGQGPMW 1872


RESULT 13
US-11-186-731-5
; Sequence 5, Application US/11186731
; Publication No. US20050255521A1
; GENERAL INFORMATION:
;  APPLICANT: Kapeller-Libermann, Rosana
;  APPLICANT:  Acton, Susan L.
;  TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
;  TITLE OF INVENTION:  Members and Uses Therefor
;  FILE REFERENCE: MPI2001-047P1RCP1(M)
;  CURRENT APPLICATION NUMBER: US/11/186,731
;  CURRENT FILING DATE:  2005-07-21
;  PRIOR APPLICATION NUMBER: US/10/077,130
;  PRIOR FILING DATE: 2002-02-15
;  PRIOR APPLICATION NUMBER: 60/269201
;  PRIOR FILING DATE: 2001-02-15
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 7968
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-186-731-5

  Query Match             35.2%;  Score 43;  DB 7;  Length 7968;
  Best Local Similarity   39.1%;  Pred. No. 7.4e+02;
  Matches    9;  Conservative    4;  Mismatches    8;  Indels    2;  Gaps    1;

Qy          2 VAEVSEFLFG--SGFAIAEGPLW 22
              ||    | ||  :|  : :||:|
Db       7188 VAGYGTFAFGGDAGGMLGQGPMW 7210


RESULT 14
US-11-087-099-11993
; Sequence 11993, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
;  APPLICANT: Abad, Mark S. et al.
;  TITLE OF INVENTION: Genes and Uses for Plant Improvement
;  FILE REFERENCE: 38-21(53450)B EP
;  CURRENT APPLICATION NUMBER: US/11/087,099
;  CURRENT FILING DATE:  2005-03-22
;  NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 11993
;   LENGTH: 461
;   TYPE: PRT
;   ORGANISM: Cytophaga hutchinsonii
US-11-087-099-11993

  Query Match             34.8%;  Score 42.5;  DB 7;  Length 461;
  Best Local Similarity   56.2%;  Pred. No. 45;
  Matches    9;  Conservative    2;  Mismatches    2;  Indels    3;  Gaps    1;

Qy         11 GSGFAIA---EGPLWT 23
              |||: ||   :| |||
Db        179 GSGYTIALKTDGTLWT 194


RESULT 15
US-11-072-512-2690
; Sequence 2690, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: 084335-0191
;  CURRENT APPLICATION NUMBER: US/11/072,512
;  CURRENT FILING DATE:  2005-03-07
;  PRIOR APPLICATION NUMBER: US 60/350,978
;  PRIOR FILING DATE: 2002-01-25
;  PRIOR APPLICATION NUMBER: JP 2001-379298
;  PRIOR FILING DATE: 2001-11-05
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2690
;   LENGTH: 444
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-072-512-2690

  Query Match             34.4%;  Score 42;  DB 7;  Length 444;
  Best Local Similarity   50.0%;  Pred. No. 52;
  Matches    7;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          9 LFGSGFAIAEGPLW 22
              | | | |  :|| |
Db         55 LLGKGLAALDGPKW 68


Search completed: March  7, 2006, 22:06:38
Job time : 10.4 secs
```

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2006, 21:53:39 ; Search time 19.2 Seconds
(without alignments)
120.271 Million cell updates/sec

Title: US-10-751-235-10
Perfect score: 122
Sequence: 1 LVAEVSEFLFGSGFAIAEGPLWTA 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	118	96.7	566	2	T46159	cytochrome P450-li
2	56	45.9	324	2	B86819	oxidoreductase ypj
3	56	45.9	344	2	G84038	phenylalanyl-tRNA
4	55	45.1	350	2	AF2572	hypothetical prote
5	54	44.3	593	2	F86441	probable cytochrom
6	52	42.6	321	2	D97836	tetraacyldisacchar
7	50	41.0	350	2	AE1227	phenylalany-tRNA s
8	50	41.0	350	2	AG1580	phenylalany-tRNA s
9	49.5	40.6	287	2	G83838	oxidoreductase (sh
10	49	40.2	287	2	C83824	ABC transporter (p
11	49	40.2	444	2	D69472	nodulation protein
12	49	40.2	494	2	S75398	lysine-tRNA ligase
13	48	39.3	465	2	AC2843	multidrug efflux p
14	48	39.3	465	2	D97620	hypothetical prote
15	48	39.3	493	2	F96696	protein F1N21.12 [
16	47	38.5	168	2	B85752	thiol peroxidase [
17	47	38.5	168	2	JC5504	thioredoxin peroxi
18	47	38.5	168	2	G90866	thiol peroxidase [
19	47	38.5	344	1	YFBSA	phenylalanine-tRNA
20	47	38.5	354	2	AB2608	phenylalanyl-tRNA
21	47	38.5	360	2	A97390	phenylalanyl-tRNA
22	47	38.5	363	2	C87336	serine proteinase
23	47	38.5	429	2	B83257	folylpolyglutamate
24	46	37.7	274	2	A25102	reaction center pr
25	46	37.7	337	2	AH0972	lipopolysaccharide
26	46	37.7	341	2	S33464	hypothetical prote
27	46	37.7	342	2	A96581	hypothetical prote
28	46	37.7	351	1	VVVP2B	coat protein VP2 -
29	46	37.7	351	1	VVVPAS	coat protein VP2 -
30	46	37.7	490	2	T31289	succinate-semialde
31	46	37.7	995	2	A88483	protein C05D11.1 [
32	45	36.9	185	2	S33294	translocon-associa
33	45	36.9	236	2	AD2820	conserved hypothet
34	45	36.9	240	2	D97598	hypothetical prote
35	45	36.9	294	2	AE1158	oxidoreductase hom
36	45	36.9	294	2	AB1517	oxidoreductase hom
37	45	36.9	344	1	VVVP2J	coat protein VP2 -
38	45	36.9	369	2	AG3502	phenylalanine-tRNA
39	45	36.9	385	2	F86464	hypothetical prote
40	45	36.9	418	2	AB0470	probable lipopolys
41	45	36.9	423	2	G96554	hypothetical prote
42	45	36.9	522	2	S13887	hydroxymethylgluta
43	45	36.9	719	2	T52043	probable glutamate
44	45	36.9	728	2	T01200	probable glutamate
45	45	36.9	819	2	T40527	hypothetical prote

ALIGNMENTS

RESULT 1
T46159
cytochrome P450-like protein - Arabidopsis thaliana
N;Alternate names: protein T4D2.60
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C;Accession: T46159
R;Nyakatura, G.; Fartmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichselgartner, M.; N
submitted to the Protein Sequence Database, December 1999
A;Reference number: Z23025
A;Accession: T46159
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-566 <NYA>
A;Cross-references: UNIPROT:Q9SCP8; UNIPARC:UPI000009D7F9; EMBL:AL132958
A;Experimental source: cultivar Columbia; BAC clone T4D2
C;Genetics:
A;Map position: 3
A;Introns: 183/2; 292/3; 358/3; 392/3; 439/3; 475/3; 503/2; 557/3
A;Note: T4D2.60
C;Superfamily: pea cytochrome P450 CYP97; cytochrome P450 homology
C;Keywords: heme; iron; metalloprotein
F;371-536/Domain: cytochrome P450 homology <P45>
F;514/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 96.7%; Score 118; DB 2; Length 566;
Best Local Similarity 100.0%; Pred. No. 1.3e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LVAEVSEFLFGSGFAIAEGPLWT 23
||||||||||||||||||||||||
Db 144 LVAEVSEFLFGSGFAIAEGPLWT 166

RESULT 2
B86819
oxidoreductase ypjF [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: B86819
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlic
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: B86819
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-324 <STO>
A;Cross-references: UNIPROT:Q9CFC7; UNIPARC:UPI00000C6A7E; GB:AE005176; PID:g12724556; PI
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: ypjF

Query Match 45.9%; Score 56; DB 2; Length 324;
Best Local Similarity 50.0%; Pred. No. 0.46;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 5 VSEFLFGSGFAIAEGPLWTA 24
Db 182 VLQFLFGTGFDAAKGIYWSS 201

RESULT 3
G84038
phenylalanyl-tRNA synthetase alpha subunit pheS [imported] - Bacillus halodurans (strain
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: G84038
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: G84038
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-344 <STO>
A;Cross-references: UNIPROT:Q9K895; UNIPARC:UPI0000136403; GB:AP001517; GB:BA000004; NID
A;Experimental source: strain C-125
C;Genetics:
A;Gene: pheS
C;Superfamily: phenylalanine-tRNA ligase alpha chain

Query Match 45.9%; Score 56; DB 2; Length 344;
Best Local Similarity 47.4%; Pred. No. 0.49;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 2 VAEVSEFLFGSGFAIAEGP 20
Db 114 ITEVEDLFIGLGFSVAEGP 132

RESULT 4
AF2572
hypothetical protein alr9005 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120ep
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AF2572
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AF2572
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-350 <KUR>
A;Cross-references: UNIPROT:Q8YJV8; UNIPARC:UPI00000CF069; GB:AP003605; PIDN:BAB77491.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr9005
A;Genome: plasmid

Query Match 45.1%; Score 55; DB 2; Length 350;
Best Local Similarity 61.1%; Pred. No. 0.72;
Matches 11; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 5 VSEFLFGSGFAIAEGPLW 22
Db 313 VDEALKAKGFVIAENPLW 330

RESULT 5
F86441
probable cytochrome P450 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: F86441
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: F86441
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-593 <STO>
A;Cross-references: UNIPROT:Q9C6S0; UNIPARC:UPI00000AA2C9; GB:AE005172; NID:g11136728; PI
C;Genetics:
A;Map position: 1
C;Superfamily: pea cytochrome P450 CYP97; cytochrome P450 homology
C;Keywords: heme; iron; metalloprotein
F;514/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 44.3%; Score 54; DB 2; Length 593;
Best Local Similarity 36.4%; Pred. No. 1.9;
Matches 8; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Qy 1 LVAEVSEFLFGSGFAIAEGPLW 22
Db 174 ILAEILDFVMGKGLIPADGEIW 195

RESULT 6
D97836
tetraacyldisaccharide 4'-kinase (EC 2.7.1.130) [imported] - Rickettsia conorii (strain Ma
C;Species: Rickettsia conorii
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: D97836
R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Rou
Science 293, 2093-2098, 2001
A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A;Reference number: A97700; MUID:21442074; PMID:11557893
A;Accession: D97836
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-321 <KUR>
A;Cross-references: UNIPROT:Q92GN1; UNIPARC:UPI000012E92C; GB:AE006914; PIDN:AAL03630.1;
C;Genetics:
A;Gene: lpxK
C;Superfamily: Escherichia coli probable tetraacyldisaccharide 4'-kinase ycaH
C;Keywords: phosphotransferase

Query Match 42.6%; Score 52; DB 2; Length 321;
Best Local Similarity 52.4%; Pred. No. 2;
Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LVAEVSEFLFGSGFAIAEGPL 21
Db 155 IVSVDSQRLFGNGFLIPAGPL 175

RESULT 7
AE1227
phenylalany-tRNA synthetase alpha chain [imported] - Listeria monocytogenes (strain EGD-
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AE1227
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AE1227
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-350 <GLA>
A;Cross-references: UNIPROT:Q8Y7Q2; UNIPARC:UPI0000055441; GB:NC_003210; PIDN:CAC99299.1
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: pheS
C;Superfamily: phenylalanine-tRNA ligase alpha chain

  Query Match 41.0%; Score 50; DB 2; Length 350;
  Best Local Similarity 40.0%; Pred. No. 4.5;
  Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 1 LVAEVSEFLFGSGFAIAEGP 20
:: |: : | |: |||||
Db 113 VIEEMEDMFIGMGYEIAEGP 132

RESULT 8
AG1580
phenylalany-tRNA synthetase alpha chain [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AG1580
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AG1580
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-350 <GLA>
A;Cross-references: UNIPROT:Q92CI7; UNIPARC:UPI0000136415; GB:AL592022; PIDN:CAC96415.1;
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: pheS
C;Superfamily: phenylalanine-tRNA ligase alpha chain

  Query Match 41.0%; Score 50; DB 2; Length 350;
  Best Local Similarity 40.0%; Pred. No. 4.5;
  Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 1 LVAEVSEFLFGSGFAIAEGP 20
:: |: : | |: |||||
Db 113 VIEEMEDMFIGMGYEIAEGP 132

RESULT 9
G83838
oxidoreductase (short chain dehydrogenase/reductase family) BH1511 [imported] - Bacillus
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 05-Oct-2004
C;Accession: G83838
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: G83838
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-287 <STO>
A;Cross-references: UNIPROT:Q9KCQ8; UNIPARC:UPI00000D73F6; GB:AP001512; GB:BA000004; NID
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH1511
C;Superfamily: short-chain dehydrogenase; short-chain alcohol dehydrogenase homology

  Query Match 40.6%; Score 49.5; DB 2; Length 287;
  Best Local Similarity 38.5%; Pred. No. 4.3;
  Matches 10; Conservative 4; Mismatches 9; Indels 3; Gaps 1;

Qy 1 LVAEVSEFLFGSGF---AIAEGPLWT 23
| :|| : | | :| ||:||
Db 202 LTRSLSESIVGQGIRVNGVAPGPIWT 227

RESULT 10
C83824
ABC transporter (permease) BH1395 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: C83824
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: C83824
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-287 <STO>
A;Cross-references: UNIPROT:Q9KD24; UNIPARC:UPI00000D73F0; GB:AP001511; GB:BA000004; NID
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH1395
C;Superfamily: conserved hypothetical protein HI0360

  Query Match 40.2%; Score 49; DB 2; Length 287;
  Best Local Similarity 40.0%; Pred. No. 5.2;
  Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 4 EVSEFLFGSGFAIAEGPLWT 23
:: :|||| |:: |||
Db 124 DLMNYLFGSVIAVSRSDLWT 143

RESULT 11
D69472
nodulation protein NfeD (nfeD) homolog - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: D69472
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.,
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: D69472
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-444 <KLE>
A;Cross-references: UNIPROT:O28493; UNIPARC:UPI0000056BB8; GB:AE000979; GB:AE000782; NID
C;Superfamily: Rhizobium nodulation competitiveness protein nfeD

  Query Match 40.2%; Score 49; DB 2; Length 444;
  Best Local Similarity 52.6%; Pred. No. 8.4;
  Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 4 EVSEFLFGSGFAIAEGPLW 22
|| || | ||| | :|
Db 381 EVLEFSNGRGFARVRGEIW 399

RESULT 12
S75398

lysine-tRNA ligase (EC 6.1.1.6) - Sulfolobus solfataricus
N;Alternate names: lysyl-tRNA synthetase; protein c04032
C;Species: Sulfolobus solfataricus
C;Date: 11-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 09-Jul-2004
C;Accession: S75398
R;Sensen, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny, S.
Mol. Microbiol. 22, 175-191, 1996
A;Title: Organizational characteristics and information content of an archaeal genome: 1
A;Reference number: S73076; MUID:97055432; PMID:8899719
A;Accession: S75398
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-494 <SEN>
A;Cross-references: UNIPROT:P95970; UNIPARC:UPI0000136543; EMBL:Y08257; NID:g1707772; PI
A;Experimental source: strain P2
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1996
C;Genetics:
A;Start codon: TTG
C;Superfamily: lysine-tRNA ligase
C;Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis

  Query Match 40.2%; Score 49; DB 2; Length 494;
  Best Local Similarity 38.1%; Pred. No. 9.5;
  Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 1 LVAEVSEFLFGSGFAIAEGPL 21
:: |: ||| || | |:
Db 175 IIREIREFLYSKGFIEVETPI 195

RESULT 13
AC2843
multidrug efflux protein [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AC2843
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AC2843
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-465 <KUR>
A;Cross-references: UNIPROT:Q8UDF5; UNIPARC:UPI00000D1DE3; GB:AE008688; PIDN:AAL43161.1;
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu2172
A;Map position: circular chromosome

  Query Match 39.3%; Score 48; DB 2; Length 465;
  Best Local Similarity 57.1%; Pred. No. 13;
  Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 8 FLFGSGFAIAEGPL 21
|:|||||::| |:
Db 71 FIFGSGFSVAVVPM 84

RESULT 14
D97620
hypothetical protein AGR_C_3941 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: D97620
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: D97620
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-465 <KUR>
A;Cross-references: UNIPROT:Q8UDF5; UNIPARC:UPI00000D1DE3; GB:AE007869; PIDN:AAK87917.1;
C;Genetics:
A;Gene: AGR_C_3941
A;Map position: circular chromosome

  Query Match 39.3%; Score 48; DB 2; Length 465;
  Best Local Similarity 57.1%; Pred. No. 13;
  Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 8 FLFGSGFAIAEGPL 21
|:|||||::| |:
Db 71 FIFGSGFSVAVVPM 84

RESULT 15
F96696
protein F1N21.12 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: F96696
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: F96696
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-493 <STO>
A;Cross-references: UNIPROT:Q9FYG3; UNIPARC:UPI000009EEE1; GB:AE005173; NID:g9828628; PI
C;Genetics:
A;Gene: F1N21.12
A;Map position: 1
C;Superfamily: glucose transport protein

  Query Match 39.3%; Score 48; DB 2; Length 493;
  Best Local Similarity 47.6%; Pred. No. 14;
  Matches 10; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

Qy 1 LVAEVSEFLFGSGFAIAEGPL 21
||| :| |||| : ||
Db 56 LVATISSFLFGYHLGVVNEPL 76

Search completed: March 7, 2006, 21:59:02
Job time : 21.2 secs